ELEMENTS OF LINEAR ALGEBRA AND MATRIX THEORY
Course by E. Kogan

Contents

PROBLEM FEATURES

Dear Student,

You will find many problems embedded into the lecture notes. I offer these problems **NOT** to check up your knowledge of the subject. I offer the problems to **TEACH** the subject. Hence, more often than not, the solution of the problem will demand from you **READING THE NOTES AND TEXTBOOKS AND STUDYING THE PROBLEMS SOLVED THERE**, until you are able to do the homework (and not just lengthy discussions with your friends). I warn you you that it will be time- and effort-consuming. But this is the most important part of studying. The reward will be the knowledge.

There will be two kinds of problems:

1. Exercises (unlabeled problems) can serve as a guide in basic learning. Those who want to know the basics of the subject should solve them all.
2. Problems (labeled with an asterisk "*") demand a deeper understanding of the subject and a bit of mathematical skill; in other words they develop a deeper understanding of the subject and a mathematical skill. Some of them can be challenging. Particularly challenging problems are identified with a double asterisk (**). It is worth trying to solve such problems only after you have solved all the other problems in the Section.

I wish you success.
Yours,
Eugene Kogan.

I. INTRODUCTION TO MATRIX ALGEBRA

A. Matrices

Definition A.1 *A matrix is a rectangular array of elements, real or complex.*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}. \quad (1)$$

(The real and complex numbers we'll call scalars.)

For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \equiv \begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix} \equiv \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad (2)$$

or

$$\begin{bmatrix} 1+6i & 2 \\ 3i & 4 \\ 5-7i & \frac{1}{4} \end{bmatrix}. \quad (3)$$

The numbers a_{ij} are called elements or entries, the subscripts i and j identify the row and the column. The matrix (1) is called a matrix of order (m, n). When $m = n$ it is called a matrix of order n or a n-square matrix.

Definition A.2 *A matrix having one column*

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} \quad (4)$$

is called an n- vector. (To save space I'll often designate such vector as $(a_1, a_2, \dots, a_n)$.)

Definition A.3 *We say that two matrices A and B are equal if they are of the same order and have equal corresponding elements for each pair of i and j:*

$$B = A \Longleftrightarrow b_{ij} = a_{ij}. \quad (5)$$

Definition A.4 *We define the sum C of two matrices A and B of the same order ($C = A + B$) as a matrix with the elements*

$$c_{ij} = a_{ij} + b_{ij}. \quad (6)$$

Problem A.1 *Given the matrices*

$$\begin{bmatrix} 1 & 5 \\ 2 & 3 \\ 4 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 5 & -2 \\ 2 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 7 & 2 & 6 \\ 2 & 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \end{bmatrix},$$

for which pairs is the sum defined? When it is defined find the sum.

Theorem A.1 *Addition is commutative and associative.*

Proof A.1 *The theorem follows directly from the commutativity and associativity of addition of scalars*

$$\begin{aligned} a_{ij} + b_{ij} = b_{ij} + a_{ij} &\Longrightarrow A + B = B + A \\ (a_{ij} + b_{ij}) + c_{ij} &= a_{ij} + (b_{ij} + c_{ij}) \\ &\Longrightarrow (A + B) + C = A + (B + C). \end{aligned} \quad (7)$$

Definition A.5 *We define the difference of two matrices A and $B (C = A - B)$ as a matrix with the elements*

$$c_{ij} = a_{ij} - b_{ij}. \quad (8)$$

Definition A.6 *We define the multiple of scalar k and matrix A ($C = kA$) as a matrix with the elements*

$$c_{ij} = ka_{ij}. \quad (9)$$

Definition A.7 *We define the product of two matrices A and B ($C = AB$) as a matrix with the elements*

$$c_{ij} = \sum_k a_{ik} b_{kj}. \quad (10)$$

The number of columns in A and the number of rows in B should be the same.

Problem A.2 *Perform the matrix multiplication*

$$a) \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 3 \end{bmatrix}; \quad b) \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \cdot \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$$

Problem A.3 *For what pairs of matrices in Exercise A.1 is the product defined in one order or the other? Find the products for these pairs.*

Problem A.4 ** True or false; give a specific counterexample when false.*

(a) If the first and third column of B are the same, so are the first and third column of AB.

(b) If the first and third rows of B are the same, so are the first and third rows of AB.

(c) If the first and third rows of A are the same, so are the first and third rows of AB.

(d) $(AB)^2 = A^2B^2$.

Theorem A.2 *Matrix multiplication is associative*

$$(AB)C = A(BC). \quad (11)$$

Proof A.2 *Applying the definition we get*

$$[(AB)C]_{ij} = \sum_l (AB)_{il} c_{lj} = \sum_{kl} a_{ik} b_{kl} c_{lj}$$
$$= \sum_k a_{ik} (BC)_{kj} = [A(BC)]_{ij}. \qquad (12)$$

Problem A.5 *Test the rule* $(AB)C = A(BC)$

$$a) \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 4 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

Matrix multiplication is in general noncommutative.

$$A \cdot B \neq B \cdot A. \qquad (13)$$

Problem A.6 *Calculate*

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

and

$$\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Check up whether the result is the same.

Problem A.7 *(a) If $AB = BA$, the matrices A and B are said to commute. If $AB = -BA$, the matrices A and B are said to anticommute.*

(a) Show that each of the matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

anticommutes with the others. These are the Pauli spin matrices, which are used in the study of electron spin in quantum mechanics.

*(b) * Show that two diagonal matrices commute.*

There are divisors of zero

$$AB = 0 \not\Longrightarrow A = 0 \text{ or } B = 0. \qquad (14)$$

Problem A.8 *Calculate*

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Theorem A.3 *Matrix multiplication is distributive with respect to addition*

$$(A + B)C = AC + BC$$
$$C(A + B) = CA + CB \qquad (15)$$

Proof A.3 *From the definition follows:*

$$[(A + B)C]_{ij} = \sum_k (a_{ik} + b_{ik}) c_{kj} = [AC + BC]_{ij}. \qquad (16)$$

The cancelation law does not hold:

$$AB = AC \not\Longrightarrow B = C. \qquad (17)$$

Problem A.9 *The sum of the main diagonal elements a_{ii}, $i = 1, 2, \ldots, n$, of a square matrix A of order n is called the trace of A:*

$$tr\,A = a_{11} + a_{22} + \ldots + a_{nn}.$$

(a) If A and B are of order n, show that

$$tr\,(A + B) = tr\,A + tr\,B.$$

(b) If C is of order (m, n) and G is of order $(n, m)B$, show that

$$tr\,CG = tr\,GC.$$

B. The Inverse of a Matrix

Definition B.1 *The matrix*

$$I_n = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix}_n \qquad (18)$$

is called the identity matrix or the unit matrix of the order n.

Definition B.2 *The inverse of a n-square matrix A (if it exists) is designated A^{-1} and must satisfy equations*

$$AA^{-1} = A^{-1}A = I_n. \qquad (19)$$

Theorem B.1 *A square matrix has at most one inverse.*

Proof B.1 *Let*

$$AB = BA = I \qquad (20)$$

and

$$AC = CA = I. \qquad (21)$$

Then

$$BA = I \to BAC = C \to B = C. \qquad (22)$$

Theorem B.2 *If the $n \times n$ matrices A and B both have inverse then*

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{23}$$

Proof B.2 *From the definition of the inverse follows*

$$B^{-1}A^{-1}AB = B^{-1}B = I. \tag{24}$$

An obvious corollary is:

$$(AB\ldots MN)^{-1} = N^{-1}M^{-1}\ldots B^{-1}A^{-1} \tag{25}$$

(if the $n \times n$ matrices $A, B, \ldots M, N$ all have inverse).

Definition B.3 *When A^{-1} exists, the matrix A is said to be invertible or nonsingular. When A^{-1} does not exist, the matrix A is said to be singular.*

Problem B.1 *Show that, if $AB = BA$ and $S^2 = B$, then also, if A^{-1} exists,*

$$(A^{-1}SA)^2 = B.$$

Problem B.2 *If $A, B,$ and C are invertible, what is the inverse of $AB^{-1}C$? Is A^2 invertible? Is $A+B$ invertible? Show that $A^{-1}(A+B)B^{-1} = A^{-1} + B^{-1}$.*

C. Polynomial Functions of Matrices

With every polynomial

$$f(x) = a_p x^p + a_{p-1}x^{p-1} + \ldots + a_1 x + a_0 \tag{26}$$

We can associate a polynomial function of an $n \times n$ matrix A

$$f(A) = a_p A^p + a_{p-1}A^{p-1} + \ldots + a_1 A + a_0 I_n. \tag{27}$$

Problem C.1 *For the following matrix A, find A^2, A^3, and $A + A^2 + A^3$:*

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

Problem C.2 ** If*

$$A = \begin{bmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{bmatrix},$$

Prove that, for all integral values of n

$$A^n = \begin{bmatrix} \cosh nx & \sinh nx \\ \sinh nx & \cosh nx \end{bmatrix}.$$

If α is a root of polynomial (26), then αI_n is the root of polynomial (27), but the equation

$$f(A) = 0 \tag{28}$$

may also have other solutions.

Problem C.3 *Check up that the equation*

$$A^2 + I_2 = 0$$

has a solution

$$A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}.$$

II. SYSTEM OF LINEAR EQUATIONS

A. Solutions for System of Equations

1. System of linear equations can have one solution:

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0. \end{aligned} \tag{29}$$

2. The system can have infinitely many solutions:

$$x_1 + x_2 = 1. \tag{30}$$

3. The system can have no solutions at all:

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned} \tag{31}$$

In the first two cases the system is called consistent, in the second - inconsistent.

B. The Row-Echelon Form for System of Equations

Consider the system

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & - & x_3 & = & -2 \\ x_1 & + & x_2 & + & x_3 & = & 0 \\ -x_1 & + & x_2 & + & 2x_3 & = & 3 \\ 3x_1 & + & 4x_2 & + & 2x_3 & = & 1 \end{array} \tag{32}$$

Let us exchange the order of equations

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ 3x_1 & + & 2x_2 & - & x_3 & = & -2 \\ -x_1 & + & x_2 & + & 2x_3 & = & 3 \\ 3x_1 & + & 4x_2 & + & 2x_3 & = & 1 \end{array} \tag{33}$$

Subtracting an equation with the appropriate coefficient from the others we consecutively get

$$\left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & -x_2 & - & 4x_3 & = & -2 \\ & & 2x_2 & + & 3x_3 & = & 3 \\ & & x_2 & - & x_3 & = & 1 \end{array}\right.$$

$$\Longrightarrow \left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & x_2 & + & 4x_3 & = & 2 \\ & & 2x_2 & + & 3x_3 & = & 3 \\ & & x_2 & - & x_3 & = & 1 \end{array}\right.$$

$$\Longrightarrow \left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & x_2 & + & 4x_3 & = & 2 \\ & & & & -5x_3 & = & -1 \\ & & & & -5x_3 & = & -1 \end{array}\right. . \tag{34}$$

We can discard the last equation and obtain

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & x_2 & + & 4x_3 & = & 2 \\ & & & & x_3 & = & 1/5 \end{array} \tag{35}$$

The system (35) is in row-echelon or Gauss form. We can easily solve it, but it is even more convenient to transform it further.

$$\left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & & & = & -\frac{1}{5} \\ & & x_2 & & & = & \frac{6}{5} \\ & & & & x_3 & = & \frac{1}{5} \end{array}\right. \Longrightarrow \left\{\begin{array}{rrrcr} x_1 & & & = & -\frac{7}{5} \\ & x_2 & & = & \frac{6}{5} \\ & & x_3 & = & \frac{1}{5} \end{array}\right. . \tag{36}$$

The method is called the sweep out process or Gauss-Jordan elimination algorithm.

Problem B.1 *Solve these systems of equations by reducing matrices to row-echelon form*

$$(a) \left\{\begin{array}{l} x_1 - x_2 + x_3 = 1 \\ 2x_1 + 3x_2 - x_3 = 4 \\ -x_1 - 2x_2 + 5x_3 = 2 \end{array}\right. \quad b) \left\{\begin{array}{l} 2x_1 + 3x_2 + 4x_3 = 8 \\ x_1 - x_2 + 2x_3 = 9 \\ -3x_1 + 2x_2 + x_3 = -4 \end{array}\right.$$

C. System of Linear Equations in Matrix Notation

Any linear system of algebraic equations

$$\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array} \tag{37}$$

We can present in the form

$$AX = B, \tag{38}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \tag{39}$$

is the matrix of coefficients,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{bmatrix}. \tag{40}$$

Equivalent to reducing to Gauss (or Gauss-Jordan) form the system of equations (37) is reducing to the same form the augmented matrix of coefficients

$$[\mathbf{A} \mid \mathbf{B}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]. \tag{41}$$

Consider, for example, the system

$$\begin{array}{r} x - 2y + z = 5 \\ 2x + y - 2z = 1 \\ 3x + y - z = 4 \end{array} \tag{42}$$

The augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & 1 & -2 & 1 \\ 3 & 1 & -1 & 4 \end{array}\right]. \tag{43}$$

Using elementary row operations, we transform the matrix $[\mathbf{A} \mid \mathbf{B}]$ into a Gauss-Jordan form:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & 1 & -2 & 1 \\ 3 & 1 & -1 & 4 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 5 & -4 & -9 \\ 0 & 7 & -4 & -11 \end{array}\right] \to$$
$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 7 & -4 & -11 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 0 & \frac{8}{5} & \frac{8}{5} \end{array}\right] \to$$
$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 0 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -2 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] \to$$
$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] \tag{44}$$

So the solution is $x = 2$, $y = -1$, $z = 1$.

Problem C.1 *Transforming the augmented matrix to the row-echelon form solve the systems*

$$\begin{bmatrix} 4 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \quad \begin{bmatrix} 3 & 5 & 1 \\ -1 & -3 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Definition C.1 *Two systems of equations are said to be equivalent if any solution of either one is also the solution of the other.*

Theorem C.1 *If matrix C is nonsingular, then equations (38) and*

$$CAX = CB \tag{45}$$

are equivalent.

Proof C.1 *It is obvious that any solution of Eq. (38) satisfies Eq. (45). But if C is invertible, it is obvious that any solution of Eq. (45) satisfies Eq.(38).*

We can easily find matrix representations of the elementary row operations. For simplicity, let us do it for the system of 3 equations.
1. Multiplying an equation number 2 by a constant k is equivalent to premultiplication of Eq. (38) by a matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{46}$$

Interchanging equations number 1 and 3 is equivalent to premultiplication of Eq. (38) by a matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{47}$$

3. Adding equation number 3 multiplied by a constant k to equation number 1 is equivalent to premultiplication of Eq. (38) by a matrix

$$\begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{48}$$

Theorem C.2 *If we apply to system (37) one of the following elementary row operations:*
1. multiplying a row by a nonzero constant;
2. interchanging two rows;
3. adding any constant multiple of one row to another row
we obtain an equivalent system of equations.

Proof C.2 *The matrix corresponding to any elementary operation is invertible:*

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/k & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 1 & 0 & -k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{49}$$

Hence the product of such matrices is also invertible.

Definition C.2 *A matrix is in row-echelon form if it satisfies the conditions:*
all nonzero rows precede zero rows;
the leftmost nonzero element of each nonzero row is unity;
the first nonzero element of any nonzero row appears in a later column than the first nonzero element of any preceding row.

D. Homogeneous System of Linear Equations

Homogeneous system of linear equations

$$AX = 0. \tag{50}$$

always has a solution $(0, 0, \ldots 0)$, which is called a trivial solution. A solution not consisting entirely of 0's is called nontrivial. The problem of interest is to determine whether or not a nontrivial solution exists, and to find it if it does. This is readily accomplished by by transforming the system into reduced echelon form. If there are n variables and if the reduced form has $r < n$ equations then we can solve for r variables in terms of the remaining $n - r$ variables.

Consider, for example the system

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 3x_1 + 2x_2 - 4x_3 &= 0 \\ 2x_1 - x_2 + x_3 &= 0. \end{aligned} \tag{51}$$

Using elementary row operations we obtain

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 3 & 2 & -4 & 0 \\ 2 & -1 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \tag{52}$$

There exists only trivial solution.

Consider another problem

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & -1 \\ 1 & 4 & -2 \\ 1 & 14 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \tag{53}$$

Using elementary row operations we obtain

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 2 & 3 & -1 & 0 \\ 1 & 4 & -2 & 0 \\ 1 & 14 & -8 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2/5 & 0 \\ 0 & 1 & -3/5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{54}$$

From this we obtain the complete solution

$$\begin{aligned} x_1 &= -2/5x_3 \\ x_2 &= 3/5x_3. \end{aligned} \tag{55}$$

Definition D.1 *A complete solution is a set of expressions which yield all the solutions of a given system.*

Generally, when we transform the homogeneous system of m linear equations of n variables to reduced echelon form, we get an equivalent system of r nontrivial equations

$$BX = 0. \tag{56}$$

If $r = n$, $B = I_n$ and the solution is $X = 0$. If $r < n$, the system can be solved for r of the variables in terms of the remaining $n - r$.

$$X = t_1X_1 + t_2X_2 + \ldots + t_{n-r}X_{n-r}, \tag{57}$$

where the vectors $X_1, \ldots, X_{n-r}$ are particular solutions. The particular solution X_i is is obtained by assigning all the t's but t_i the value 0 and assigning t_i the value 1.

For example, consider

$$\begin{aligned} x_1 - 2x_2 + x_3 + 4x_4 - x_5 &= 0 \\ 2x_1 + x_2 - x_3 + 5x_4 + x_5 &= 0 \\ x_1 + 13x_2 - 8x_3 - 5x_4 + 8x_5 &= 0. \end{aligned} \tag{58}$$

The reduced echelon form of the augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1/5 & 14/5 & 1/5 & 0 \\ 0 & 1 & -3/5 & -3/5 & 3/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{59}$$

Now we put $x_3 = 5t_1$, $x_4 = 5t_2$, $x_3 = 5t_3$, and get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = t_1 \begin{bmatrix} 1 \\ 3 \\ 5 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -14 \\ 3 \\ 0 \\ 5 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -1 \\ -3 \\ 0 \\ 0 \\ 5 \end{bmatrix}. \tag{60}$$

Problem D.1 *Obtain complete solutions for the following systems of equations:*
a)

$$\begin{aligned} x_1 - x_2 + x_3 + x_4 + 2 &= 0 \\ x_1 + x_2 - x_3 + 2x_4 - 1 &= 0 \\ 3x_1 - x_2 + x_3 + 2x_4 - 2 &= 0 \end{aligned}$$

b)

$$\begin{aligned} x_1 - x_2 + x_3 + x_4 + 2 &= 0 \\ x_1 + x_2 - x_3 + 2x_4 - 1 &= 0 \\ 3x_1 - x_2 + x_3 + 4x_4 + 4 &= 0 \end{aligned}$$

Problem D.2 *For what values of k will the equation*

$$\begin{bmatrix} 1 & 1 & k \\ 1 & k & 1 \\ k & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

have nontrivial solutions? What are these solutions in each case?

Problem D.3 * *Express the solutions of the system*

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + \alpha x_4 &= 0 \\ x_1 + \alpha x_2 + 3x_3 + 2x_4 &= 0 \\ \alpha x_1 + x_2 + 3x_3 + x_4 &= 0 \end{aligned}$$

as functions of the parameter α.

E. Nonhomogeneous System of Linear Equations

Consider nonhomogeneous systems of linear equations

$$AX = B. \tag{61}$$

For example, consider the system

$$\begin{aligned} x - 2y + z &= 5 \\ 2x + y - 2z &= 1 \\ 3x - y - z &= 3. \end{aligned} \tag{62}$$

Using the sweeping algorithm we get

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & 1 & -2 & 1 \\ 3 & -1 & -1 & 3 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 5 & -4 & -9 \\ 0 & 5 & -4 & -12 \end{array}\right] \rightarrow$$
$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 5 & -4 & -12 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 0 & 0 & -3 \end{array}\right]. \tag{63}$$

The last equation of the system is

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 0. \tag{64}$$

Hence the system is inconsistent.

Consider the system

$$\begin{aligned} x - 2y + z &= 5 \\ 2x + y - 2z &= 1 \\ 3x - y - z &= 6 \end{aligned} \tag{65}$$

Using the algorithm we get

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & 1 & -2 & 1 \\ 3 & -1 & -1 & 6 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 5 & -4 & -9 \\ 0 & 5 & -4 & -9 \end{array}\right] \rightarrow$$
$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 5 & -4 & -9 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -\frac{4}{5} & -\frac{9}{5} \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{66}$$

Hence the third equation can be discarded and the system can be rewritten in the form

$$x - 2y + z = 5$$
$$y - \frac{4}{5}z = -\frac{9}{5}. \tag{67}$$

We solve the system only for the leftmost variables. Any other variables are arbitrary. So the system (67) can be presented by the augmented matrix

$$\left[\begin{array}{cc|c} 1 & -2 & \frac{7}{5} - \frac{3}{5}z \\ 0 & 1 & -\frac{9}{5} + \frac{4}{5}z \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 5 - z \\ 0 & 1 & -\frac{9}{5} + \frac{4}{5}z \end{array}\right], \tag{68}$$

with the solution

$$x = \frac{7}{5} - \frac{3}{5}z$$
$$y = -\frac{9}{5} + \frac{4}{5}z. \tag{69}$$

Consider the system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 10 \\ 4x_1 + 5x_2 + 6x_3 &= 16 \\ 7x_1 + 8x_2 + 9x_3 &= \alpha \end{aligned} \tag{70}$$

To solve we can transform the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 4 & 5 & 6 & 16 \\ 7 & 8 & 9 & \alpha \end{array}\right] \tag{71}$$

into a row echelon form:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 4 & 5 & 6 & 16 \\ 7 & 8 & 9 & \alpha \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 0 & -3 & -6 & -24 \\ 0 & -6 & -12 & \alpha - 70 \end{array}\right]$$
$$\rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & -6 & -12 & \alpha - 70 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & \alpha - 22 \end{array}\right]. \tag{72}$$

The system can be solved only if $\alpha = 22$. In this case any equation can be represented as a linear combination of two other equations.(Multiplying second equation by two and distracting from the product the first equation we get the third equation.) So we can throw any equation, and solve the resulting system of two equations.

Theorem E.1 *The complete solution of a nonhomogeneous system $AX = B$ is the sum of any particular solution of the given system and the complete solution of the corresponding homogeneous system $AX = 0$.*

Proof E.1

For example, consider

$$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= 0 \\ 2x_1 - x_2 + x_3 - x_4 &= 3 \\ x_1 + 4x_2 - 10x_3 + 4x_4 &= -3. \end{aligned} \tag{73}$$

The reduced echelon form is

$$\left[\begin{array}{cccc|c} 1 & 0 & -\frac{2}{3} & 0 & 1 \\ 0 & 1 & -\frac{7}{3} & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{74}$$

The solution in vector notation is

$$X = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} 2 \\ 7 \\ 3 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}. \tag{75}$$

Problem E.1 *Solve the system of equations*

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ x_1 + x_2 - x_3 - x_4 &= 3 \\ x_1 - x_2 - x_3 - x_4 &= 4 \end{aligned}$$

Problem E.2 *Under what conditions on t will these systems be consistent? What are the solutions in these cases?*
a)

$$\begin{aligned} 3x_1 + x_2 + x_3 &= t \\ x_1 - x_2 + 2x_3 &= 1 - t \\ x_1 + 3x_2 - 3x_3 &= 1 + t, \end{aligned}$$

b)

$$\begin{aligned} tx_1 + x_2 &= 0 \\ x_1 + tx_2 - x_3 &= 1 \\ -x_2 + tx_3 &= 0. \end{aligned}$$

F. Computation of the Inverse of a Matrix

Theorem F.1 *To compute the inverse of a square matrix A we can use the following algorithm:*
Step 1: Form the partitioned matrix $[A|I]$, where I is the identity matrix having the same order as A.
Step 2: Using elementary row operations, transform A into row-echelon form, applying each operation, to the entire matrix formed in Step 1. Denote the result as $[C|D]$.
Step 3: If C has a zero row, stop; the original matrix is singular. Otherwise continue.
Step 4: Beginning with the last column of C and progressing backward iteratively through the second column, use elementary row operations to transform all elements above the diagonal of C to zero. Apply each operation, however, to the entire matrix $[C|D]$. Denote the result as $[I|B]$. The matrix B is the inverse of the original matrix A.

Let us invert, for example, the matrix

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}. \tag{76}$$

We obtain

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right] \tag{77}$$

Hence

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \tag{78}$$

Proof F.1 *Consider system of linear equations*

$$AX = B, \tag{79}$$

where A is a non-singular matrix. The solution has the form

$$X = A^{-1}B. \tag{80}$$

Hence, when we apply the Gauss-Jordan algorithm to the partitioned matrix $[A|B]$ we get

$$[I_n|A^{-1}B]. \tag{81}$$

More generally, one can treat the system of equations

$$AX = B_1,\ AX = B_2, \dots, AX = B_p \tag{82}$$

simultaneously, by applying the algorithm to the array

$$[A|B_1, B_2, \dots B_p]. \tag{83}$$

We get

$$[I_n|A^{-1}B_1, A^{-1}B_2, \dots A^{-1}B_p]. \tag{84}$$

If we chose B_js to be elementary n-vectors $E_1, E_2 \dots E_n$ (where E_j is an n-vector with 1 in the jth row, all other elements being zero), then the initial array is

$$[A|B_1, B_2, \dots B_p] = [A|I_n], \tag{85}$$

and the final array is

$$[I_n|A^{-1}B_1, A^{-1}B_2, \dots A^{-1}B_p] = [I_n|A^{-1}]. \tag{86}$$

Problem F.1 *Find the inverse of the matrix*

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}.$$

Problem F.2 *Find the inverse of the given matrix if it exists.*

$$(a) \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 3 & 2 & 5 \end{bmatrix} \quad (b) \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 2 \\ 3 & -1 & -1 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ -1 & 0 & 4 \end{bmatrix}$$
$$(d) \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} \quad (e) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (f) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$(g) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (h) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

III. DETERMINANTS

A. The Definition of a Determinant

Consider a system of two linear equations with 2 unknowns

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2. \end{aligned} \tag{87}$$

The solution is

$$x = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}, \quad y = \frac{a_{12}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}. \tag{88}$$

Let us for a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{89}$$

define a function (which we shall call determinant) by the relation

$$\det A = ad - bc. \tag{90}$$

Problem A.1 *Let*

$$A = \begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix} \quad B = \begin{bmatrix} 8 & -3 \\ 5 & -6 \end{bmatrix}.$$

Calculate $\det A$, $\det B$ *and* $\det AB$.

Definition A.1 *A permutation π of the integers $1, 2, \dots, n$ is a linear arrangement of the integers in some order.*

For example, two possible permutations of 1 and 2 are

$$\pi_1 = \{1, 2\} \quad \text{and} \quad \pi_2 = \{2, 1\}. \tag{91}$$

Definition A.2 *If for each of the integers in permutation we count the number of smaller integers, following it, and add the results, we get the parity index of the permutation $\mu(\pi)$.*

Definition A.3 *The determinant of the $n \times n$ matrix A*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{92}$$

is designated det A or $|A|$ and is defined

$$det\ A = \sum_{\pi} (-1)^{\mu(\pi)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}, \tag{93}$$

where the summation extends over all $n!$ permutations π.

Problem A.2 *Calculate*

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

using the definition (93).

The calculation of determinant based on the definition is a very uneconomical procedure. In fact,the number of terms in Eq. (93) is $n!$. If we have matrix 20×20 the number of terms in Eq. (93) is $20! \approx 2.4 \times 10^{18}$. So, no existing computer can calculate the determinant that way. To develop more efficient algorithms, we should study the properties of determinants.

Problem A.3 *Calculate*

$$(a)\ \det \begin{vmatrix} 1 & -2 & 3 \\ 3 & 5 & 1 \\ 6 & 4 & 2 \end{vmatrix};\ (b)\ \det \begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & -1 \\ 2 & -1 & 1 \end{vmatrix}.$$

Problem A.4 *Show that*

$$\begin{vmatrix} \cos\theta & 1 & 0 \\ 1 & 2\cos\theta & 1 \\ 0 & 1 & 2\cos\theta \end{vmatrix} = \cos 3\theta$$

B. Properties of Determinants

Theorem B.1 *If all elements of any row or column of a matrix A are 0, then $|A| = 0$.*

Proof B.1 *Every row and every column of A has an entry in each term of the determinant.*

Theorem B.2 *There exists a column expansion of a determinant (93):*

$$det\ A = \sum_{\pi} (-1)^{\mu(\pi)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}. \tag{94}$$

Proof B.2 *Each row and each column is represented exactly once in each product both in (93) and in (94). Consider the product $a_{1j_1} a_{2j_2} \cdots a_{nj_n}$. Move the factor a_{kn} to the far right. The parity index of the permutation of column indices is reduced by $n-k$ and the parity index of the permutation of row indices is increased by $n-k$. So the difference of the parity indices changed by an even number. But using such operations we can obtain the permutation $a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$. Hence the difference of the parity indices of the permutation of the column indices $a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ and the permutation of the row indices $a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$ is an even number. Hence a given term has the same sign in row and column expansions.*

Theorem B.3 *If two rows (columns) of a matrix A are interchanged, the sign of the determinant is changed.*

Proof B.3 *Each term of the row expansion of det A appears once in the row expansion of the determinant of the matrix with the interchanged rows, but with the opposite sign.*

Theorem B.4 *If two rows (columns) of a matrix A are identical, then $|A| = 0$.*

Proof B.4 *If we interchange two identical lines we do not change the matrix. Hence,*

$$det\ A = -det\ A, \tag{95}$$

and we get det $A = 0$.

Theorem B.5 *If we add to a line of a matrix any multiple of another line, the determinant does not change.*

Proof B.5 *The proof is based on the equation*

$$\begin{aligned} &det[A_1, \dots, A_{i-1}, A_i + A_k, A_{i+1}, \dots, A_n] \\ &= det[A_1, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_n] \\ &+ det[A_1, \dots, A_{i-1}, A_k, A_{i+1}, \dots, A_n] \\ &= det[A_1, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_n]. \end{aligned} \tag{96}$$

Theorem B.6 *The determinant of a matrix in an echelon form is the product of diagonal elements.*

Proof B.6

These properties is a key to the algorithm of calculating the determinant of a matrix based on reducing it to a Gaussian form.

For example,

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{vmatrix}$$

$$= -3 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -11 \end{vmatrix} = -3 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = -3. \qquad (97)$$

Using this algorithm, a determinant 20×20 for example, can be calculated by hand.

Problem B.1 *Prove that, if A is skew-symmetric and of odd order, then det $A = 0$.*

Problem B.2 *If every row of A adds up to zero, prove that det $A = 0$.*

Problem B.3 *Show that, in a square matrix A all the elements for which the sum of the subscripts is odd are multiplied by -1, the determinant of the new matrix equals det A.*

C. Expansion by Cofactors

Chose an arbitrary row i of the $n \times n$ matrix A. Any term in the definition of determinant (92) contains one and only one element from the row i. Hence the det A can be expanded in terms of the element of the ith row:

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij}, \qquad (98)$$

where the cofactors A_{ij} do not depend upon the elements of the row i. The expansion of det A in terms of the element of the kth column:

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij} \qquad (99)$$

Theorem C.1 *In order to determine the cofactor A_{ij} of a_{ij} in the matrix $[a_{ij}]$, delete the ith row and jth column of A, compute the determinant (called a "minor") of the remaining "sub-matrix", and multiply the result by $(-1)^{i+j}$.*

Proof C.1 *If we group all the terms in det A which contain a_{nn} and factor out a_{nn} we get*

$$A_{nn} a_{nn} = \Big(\sum (-1)^{\mu(j_1,\ldots,j_{n-1},n)} \cdot a_{1j_1} a_{2j_2} \cdots a_{n-1 j_{n-1}} \Big) a_{nn}. \qquad (100)$$

The sum in parentheses is the expansion of

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n-1} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n-1} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n-1} \end{vmatrix}. \qquad (101)$$

We also can move any entry a_{ij} to the n,n position by interchanging columns and rows $2n - i - j$ times. This proves the theorem.

This theorem can serve as a basis for the recurrent algorithm for the calculation of det A. The simple illustration of the algorithm follows:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 0, \quad (102)$$

where we expanded the determinant with respect to elements of the first row. Expanding with respect to elements of the second column we get

$$\det A = -2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 5 \cdot \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} - 8 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 0. \quad (103)$$

The row expansion can be useful if we have a lot of zeros in a row. For example, let us compute the determinant of the tridiagonal matrix

$$I_n = \begin{vmatrix} a & a & 0 & \ldots & \ldots & \ldots & 0 \\ b & a & a & 0 & \ldots & \ldots & 0 \\ 0 & b & a & a & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 0 & b & a & a \\ 0 & \ldots & \ldots & \ldots & \ldots & b & a \end{vmatrix}_n. \qquad (104)$$

(A tridiagonal matrix is one in which nonzero entries may appear only on the main diagonal and the two adjacent diagonals.) Using the row expansion twice we get the recurrent relation

$$I_n = aI_{n-1} - abI_{n-2}. \qquad (105)$$

Eq. (105) is an example of difference equation. The most famous difference equation is due to Fibonacci

$$F_{k+2} = F_{k+1} + F_k. \qquad (106)$$

It is supplemented by initial conditions

$$F_0 = 1, \quad F_1 = 1. \qquad (107)$$

Looking for a solution in the form

$$F_k = \lambda^k, \qquad (108)$$

we obtain

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}. \qquad (109)$$

Using the initial conditions we get

$$F_k = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{k+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{k+1} \right]. \qquad (110)$$

Now let us return to Eq. (105). Looking for a solution in the form (108) we obtain

$$\lambda_\pm = \frac{a \pm \sqrt{a^2 - 4ab}}{2}. \qquad (111)$$

Using the initial conditions

$$\begin{aligned} I_1 &= a \\ I_2 &= a(a-b), \end{aligned} \qquad (112)$$

we get

$$I_n = \frac{1}{\sqrt{a^2 - 4ab}} \left[\lambda_+^{n+1} - \lambda_-^{n+1} \right]. \qquad (113)$$

Problem C.1 *Show that the value of the determinant*

$$\begin{vmatrix} 1-n & 1 & 1 & \dots & 1 \\ 1 & 1-n & 1 & \dots & 1 \\ 1 & 1 & 1-n & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1-n \end{vmatrix}_n$$

is zero.

Problem C.2 ** Prove that*

$$\begin{vmatrix} x+\lambda & x & x & \dots & x \\ x & x+\lambda & x & \dots & x \\ x & x & x+\lambda & \dots & x \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & x+\lambda \end{vmatrix}_n = \lambda^{n-1}(nx+\lambda).$$

Problem C.3 *** Prove, using mathematical induction, that if*

$$V = \begin{bmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{bmatrix},$$

then

$$det\, V = \prod_{1 \le i < j \le n} (x_i - x_j).$$

The matrix V is called a **Vandermonde matrix**.

Problem C.4 *** This result appeared in a research paper on economics:*

$$\begin{vmatrix} \frac{1+a_1}{a_1} & 1 & 1 & \dots & 1 \\ 1 & \frac{1+a_2}{a_2} & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & \frac{1+a_n}{a_n} \end{vmatrix} = \frac{1 + \sum a_i}{\prod a_i}.$$

Prove that the result is correct.

Problem C.5 *** Compute*

$$\begin{vmatrix} a & a & \dots & a & a \\ b & a & \dots & a & a \\ 0 & b & \dots & a & a \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a & a \\ 0 & 0 & \dots & b & a \end{vmatrix}_n.$$

Problem C.6 *** Evaluate the Cauchy determinant*

$$\left| \frac{1}{\lambda_i + \mu_j} \right| \qquad i,j = 1,2,\dots,N$$

D. The Transpose of a Matrix

Definition D.1 *The transpose of matrix A is defined*

$$[A^T]_{ij} = [A]_{ji} \qquad (114)$$

For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \qquad (115)$$

Theorem D.1 *The transpose of the product is equal to the written in reverse order product of transposed. reverse order*

$$(AB)^T = B^T A^T \qquad (116)$$

Proof D.1 *From the definition of the transpose follows*

$$\begin{aligned} [(AB)^T]_{ij} &= [(AB)]_{ji} = \sum_k a_{jk} b_{ki} \\ &= \sum_k [A^T]_{kj} [B^T]_{ik} = [B^T A^T]_{ij} \end{aligned} \qquad (117)$$

Definition D.2 *A matrix is called symmetric, if*

$$A = A^T \Longleftrightarrow a_{ij} = a_{ji}; \qquad (118)$$

A matrix is called skew-symmetric, if

$$A = -A^T \Longleftrightarrow a_{ij} = -a_{ji}. \qquad (119)$$

Problem D.1 *Which of the two matrix is symmetric, and which is skew-symmetric*

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}; \quad \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$$

is skew-symmetric?

Theorem D.2 *Transposition of the matrix leaves the determinant invariant.*

$$det\, A^T = det\, A \qquad (120)$$

Proof D.2 *The row expansion of det A^T is precisely the column expansion of det A. It means, that in every theorem about determinants, it is legitimate to interchange the words "row" and "column" throughout.*

E. Matrix Inversion by Determinants

Theorem E.1 *If det $A \neq 0$, A^{-1} exists and is given by the equation*

$$A^{-1} = \frac{1}{det\ A} \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}, \tag{121}$$

where A_{ij}'s are the cofactors of the a_{ij}'s in det A. (The matrix $[A_{ij}]^T$ is called the adjoint matrix of A.)

Proof E.1 *The expansions (98) implies that*

$$\sum_{j=1}^{n} c_j A_{ij} \tag{122}$$

is the determinant of a matrix the same as A except that the ith row of A has been replaced by c's. Hence, if we multiply the elements of a line on the cofactors of the elements of the other line we obtain 0, because it is determinant of a matrix with two identical rows. Combining it with Eq. (98), we obtain

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \delta_{ik}\ det\ A, \tag{123}$$

which proves the theorem.

For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \frac{1}{\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}} \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}^T = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}. \tag{124}$$

Eq. (121), being concise and beautiful, is a very uneconomical way to inverse a large matrix, because it demands calculation of a lot of determinants.

Problem E.1 *Which of the following matrices have inverses? Justify your answers and find the inverses that exist.*

$$(a) \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, (b) \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

Problem E.2 *Find by determinants the inverse of a given matrix if it exists*

$$(a) \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 3 & 2 & 5 \end{bmatrix}, (b) \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}, (c) \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix}.$$

Problem E.3 *Given that $a^3 + b^3 = 1$, find by determinants the inverse of*

$$\begin{bmatrix} a & b & 0 \\ 0 & a & b \\ b & 0 & a \end{bmatrix}.$$

Consider a system

$$AX = B. \tag{125}$$

Suppose that A^{-1} exists. Then

$$X = A^{-1}B, \tag{126}$$

or in scalar form

$$x_k = \frac{\sum_{i=1}^{n} A_{ik} b_i}{|A|} \tag{127}$$

This is called Cramer's rule. An obvious corollary is

Theorem E.2 *A system of n linear homogeneous equations in n unknowns has only the trivial solution if $|A| \neq 0$.*

Proof E.2 *Let $|A| \neq 0$. Then A^{-1} exists, and*

$$AX = 0 \longrightarrow X = A^{-1}0 = 0. \tag{128}$$

Problem E.4 *Solve for x, y, z by Cramer's rule*

$$\begin{cases} x + 2y + z = 4 \\ x - y + z = 5 \\ 2x + 3y - z = 1 \end{cases}$$

F. Determinants as Multilinear Functionals

The theorems of Subsection III B may be stated in a combined form

Theorem F.1 *If we define the function D of n $n-$ vectors by the relation*

$$D(X_1, X_2, \dots, X_n) = detA, \tag{129}$$

where A is the matrix whose ordered rows are $X_1, X_2, \dots, X_n$, then:
1. D is multilinear functional, that is

$$\begin{aligned} &D(X_1, \dots, X_{i-1}, cX_i + d\tilde{X}_i, X_{i+1}, \dots, X_n) \\ &= cD(X_1, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_n) + \\ &\quad dD(X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_n). \end{aligned} \tag{130}$$

2. $D(X_1, X_2, \dots, X_n)$ is altered only in algebraic sign if two of the vectors $X_1, X_2, \dots, X_n$ are interchanged. (Hence, if any two vectors X_j and X_k are equal, the functional is equal to zero.)
3. $D(E_1, E_2, \dots E_n) = 1$.

Proof F.1 *The proof is contained in the Subsection III B.*

Theorem F.2 *If a function $f(X_1, X_2, \ldots, X_n)$ of n $n-$vectors has Properties 1 and 2, then*

$$f(X_1, X_2, \ldots, X_n) = det\, A\; f(E_1, E_2, \ldots, E_n), \quad (131)$$

where A is the matrix whose ordered rows (columns) are $X_1, X_2, \ldots, X_n$.

Proof F.2 *Let us present X_i as*

$$X_i = a_{ij} E_j. \quad (132)$$

Then for the function D we obtain

$$\begin{aligned} & f(X_1, X_2, \ldots, X_n) = f(a_{1j_1}E_{j_1}, a_{2j_2}E_{j_2}, \ldots, a_{nj_n}E_{j_n}) \\ & = a_{1j_1}a_{2j_2}\ldots a_{nj_n} f(E_{j_1}, E_{j_2}, \ldots, E_{j_n}) \\ & = \sum_{\pi} (-1)^{\mu(\pi)} a_{1j_1}a_{2j_2}\ldots a_{nj_n} f(E_1, E_2, \ldots, E_n) \\ & = det\, A\; f(E_1, E_2, \ldots, E_n). \end{aligned} \quad (133)$$

Theorem F.3 *The determinant of the product of two matrices is the product of their determinants.*

$$det\,(AB) = det\, A\; det\, B. \quad (134)$$

Proof F.3 *Let A is an arbitrary $n\times n$ matrices. We now define the function f so that*

$$f(X_1, X_2, \ldots, X_n) = D(AX_1, AX_2, \ldots, AX_n), \quad (135)$$

noting that f inherits Properties 1 and 3. Hence,

$$f(X_1, \ldots, X_n) = D(X_1, \ldots, X_n) f(E_1, \ldots, E_n), \quad (136)$$

or

$$D(AX_1, \ldots, AX_n) = D(X_1, \ldots, X_n) D(AE_1, \ldots, AE_n). \quad (137)$$

If we chose the columns of an arbitrary matrix B as $\{X_i\}$, we obtain Eq. (134).

Theorem F.4 *A^{-1} exists only if $det\, A \neq 0$.*

Proof F.4 *Suppose that A^{-1} exists. Then*

$$det\, A \cdot det\, A^{-1} = 1. \quad (138)$$

Hence $det\, A \neq 0$.

Problem F.1 ** If*

$$A = [A_1, A_2, \ldots, A_m],$$

where the A_j are n-vectors, find representation of A^T and $A^\dagger$ in terms of the A_j's.

Problem F.2 *** Calculate the determinant*

$$det\, B = \begin{vmatrix} 1+a_1 & a_1 & a_1 & \ldots & a_1 \\ a_2 & 1+a_2 & a_2 & \ldots & a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_n & 1 & a_n & \ldots & 1+a_n \end{vmatrix}.$$

Hint: If you introduce vector $A = (a_1, a_2, \ldots, a_n)^T$, then the determinant can be presented as

$$det\, B = D(E_1 + A, E_2 + A, \ldots, E_n + A).$$

Calculate the determinant using properties of multilinear functionals.

IV. VECTOR SPACES

A. Geometric Vectors

Definition A.1 *If AB and CD are directed line segments, they are said to define the same geometric vector - also designated as $\overrightarrow{AB}$ or $\overrightarrow{CD}$ - provided the segments have the same length and direction. Alternatively we designate geometric vector by a single boldface letter.*

Definition A.2 *If $\mathbf{a}$ and $\mathbf{b}$ are geometric vectors, to determine the sum $\mathbf{a}+\mathbf{b}$: place the tail of a segment representing b on the head of a segment representing $\mathbf{a}$, and join the tail of the latter to the head of the former (or use the parallelogram rule).*

Definition A.3 *If $\mathbf{a}$ is a geometric vector, then $r\mathbf{a}$ is a geometric vector that may be represented by a segment $|r|$ times as long and having the same or opposite direction according as $r > 0$ or $r < 0$.*

Problem A.1 *Represent two geometric vector α and β by directed line segments drawn at random in the plane, and then construct a representation of $r\alpha + s\beta$, where (a) $r = 1/2, s = 2$; (b) $r = -2, s = 3$.*

Problem A.2 *Use vector methods to determine whether the following points in ordinary space are vertices of parallelogram: $(5, 5, 1)$, $(3, 7, -2)$, $(7, 6, 0)$, $(5, 8, -3)$.*

Definition A.4 *The scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is*

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta. \quad (139)$$

Definition A.5 *The vector product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector in a direction perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ with the length given by the equation*

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin\theta. \quad (140)$$

Theorem B.1 *The equation of the line passing through a point* $\mathbf{a}$ *and having a direction* $\mathbf{b}$ *is*

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}, \tag{141}$$

or, alternatively

$$(\mathbf{r} - \mathbf{a})\times\mathbf{b} = 0. \tag{142}$$

or (in a coordinate form)

$$\frac{x - a_x}{b_x} = \frac{y - a_y}{b_y} = \frac{z - a_z}{b_z}. \tag{143}$$

Proof B.1

Theorem B.2 *The equation of the plane passing through a point* $\mathbf{a}$ *and perpendicular to a unit vector* $\mathbf{n}$ *is*

$$(\mathbf{r} - \mathbf{a})\cdot\mathbf{n} = 0. \tag{144}$$

Proof B.2

Theorem B.3 *The equation of the plane passing through the points* $\mathbf{a}, \mathbf{b}$ *and* $\mathbf{c}$ *is*

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) + \mu(\mathbf{c} - \mathbf{a}), \tag{145}$$

or (in a coordinate form)

$$lx + my + nz = d. \tag{146}$$

Proof B.3

1. Using vectors to find distances

Theorem B.4 *The distance from a point* $\mathbf{p}$ *to a line* $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$ *is*

$$d = |(\mathbf{p} - \mathbf{a}) - ((\mathbf{p} - \mathbf{a})\cdot\hat{\mathbf{b}})\hat{\mathbf{b}}| = |(\mathbf{p} - \mathbf{a})\times\hat{\mathbf{b}}|, \tag{147}$$

where $\hat{\mathbf{b}} = \frac{\mathbf{b}}{|\mathbf{b}|}$.

Proof B.4 *The distance from a point* $\mathbf{p}$ *to an arbitrary point on the line is*

$$d_\lambda = |\mathbf{p} - \mathbf{a} - \lambda\mathbf{b}|. \tag{148}$$

The distance from a point $\mathbf{p}$ *to a line is the minimum of* d_λ *with respect to* λ. *Hence*

$$\lambda_{min}\mathbf{b}^2 - (\mathbf{p} - \mathbf{a})\cdot\mathbf{b} = 0. \tag{149}$$

Problem B.1 *Find the distance from the point* P *with coordinates* $(1, 2, 1)$ *to the line* $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$ *where* $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$ *and* $\mathbf{b} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$.

Theorem B.5 *The distance from a point* $\mathbf{p}$ *to a plane* $(\mathbf{r} - \mathbf{a})\cdot\mathbf{n} = 0$ *is*

$$d = |(\mathbf{p} - \mathbf{a})\cdot\mathbf{n}|. \tag{150}$$

Proof B.5 *To find the distance from a point to a plane we should minimize with respect to* $\mathbf{r}$ *the distance* $d_r = |\mathbf{p} - \mathbf{r}|$ *(it will be more convenient to minimize* d_r^2*). Using the method of Lagrange multipliers, we can look for the unconditional minimum of*

$$d_r^2 - \lambda(\mathbf{r} - \mathbf{a})\cdot\mathbf{n}, \tag{151}$$

with the solution

$$\mathbf{r}_{min} = \mathbf{p} - \frac{\lambda}{2}\mathbf{n}. \tag{152}$$

The value of λ *is then found from the condition, that the point* $\mathbf{r}_{min}$ *should belong to the plane.*

Problem B.2 *Find the distance from the point* P *with coordinates* $(1, 2, 3)$ *to the plane which contains the points* A, B *and* C *with coordinates* $(0, 1, 0), (2, 3, 1)$ *and* $(5, 7, 2)$.

Theorem B.6 *The distance from the line* $\mathbf{r} = \mathbf{b}_1 + \mathbf{a}_1$ *to the line* $\mathbf{r} = \mathbf{b}_2 + \mathbf{a}_2$ *is*

$$d = |(\mathbf{a}_1 - \mathbf{a}_2)\cdot\mathbf{n}|, \tag{153}$$

where

$$\mathbf{n} = \frac{\mathbf{b}_1\times\mathbf{b}_2}{|\mathbf{b}_1\times\mathbf{b}_2|}. \tag{154}$$

Proof B.6

Problem B.3 *A line is inclined at equal angles to the* $x-$ $y-$ *and* $z-$ *axes, and passes through the origin. Another line passes through the points* $(1, 2, 4)$ *and* $(0, 0, 1)$. *Find the distance between the two lines.*

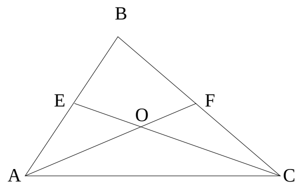
Theorem B.7 *The distance from the line* $\mathbf{r} = \mathbf{b} + \mathbf{a}_1$ *to the plane* $(\mathbf{r} - \mathbf{a}_2)\cdot\mathbf{n} = 0$ *is*

$$d = |(\mathbf{a}_1 - \mathbf{a}_2)\cdot\mathbf{n}|. \tag{155}$$

Proof B.7

Theorem B.8 *The medians of a triangle intersect at a point located two-thirds of the distance along any median drawn from its associated vertex.*

Proof B.8 *From the figure.*



we see

$$\begin{aligned}\overrightarrow{AO} &= r\overrightarrow{AF}\\ \overrightarrow{EO} &= s\overrightarrow{EC}\\ \overrightarrow{AO} &= \frac{1}{2}\overrightarrow{AB} + \overrightarrow{EO}\\ \overrightarrow{AF} &= \overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC}\\ \overrightarrow{EC} &= \frac{1}{2}\overrightarrow{AB} + \overrightarrow{BC}.\end{aligned} \tag{156}$$

Hence

$$r(\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC}) = \frac{1}{2}\overrightarrow{AB} + s(\frac{1}{2}\overrightarrow{AB} + \overrightarrow{BC}), \tag{157}$$

or equivalently

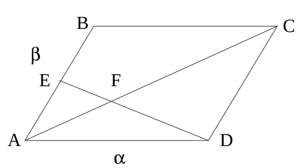
$$(\frac{1}{2} + \frac{s}{2} - r)\overrightarrow{AB} + (s - \frac{r}{2})\overrightarrow{BC} = 0. \tag{158}$$

So we obtain

$$\left\{\begin{array}{c}\frac{1}{2} + \frac{s}{2} - r = 0\\ s - \frac{r}{2} = 0\end{array}\right. \Longrightarrow \left\{\begin{array}{c} s = 1/3\\ r = 2/3\end{array}\right. . \tag{159}$$

Theorem B.9 *The line segment joining one vertex of a parallelogram with the midpoint of the opposite side trisects the intersecting diagonal and is itself trisected.*

Proof B.9 *From the figure.*



we see

$$\begin{aligned}\overrightarrow{AF} &= s\overrightarrow{AC} = s(\overrightarrow{AD} + \overrightarrow{DC})\\ \overrightarrow{EF} &= r\overrightarrow{ED} = r(\overrightarrow{AD} - \overrightarrow{AE})\\ \overrightarrow{AE} &= \frac{1}{2}\overrightarrow{DC}\\ \overrightarrow{AF} &= \overrightarrow{AE} + \overrightarrow{EF}.\end{aligned} \tag{160}$$

Hence

$$(s - r)\overrightarrow{AE} + (s - \frac{1}{2} + \frac{r}{2})\overrightarrow{DC} = 0. \tag{161}$$

So we obtain

$$\left\{\begin{array}{c} s - r = 0\\ s - \frac{1}{2} + \frac{r}{2} = 0\end{array}\right. \Longrightarrow \left\{\begin{array}{c} s = 1/3\\ r = 1/3\end{array}\right. . \tag{162}$$

Problem B.4 *Use vector methods to establish each of the following prepositions:*
(a) The diagonals of the parallelogram bisect each other;
(b) The line segment joining the midpoints of two sides of a triangle is parallel to and half as long as the third side;
(c) In any quadrilateral the lines joining the midpoints of opposite sides bisect each other.

Problem B.5 * *Use vector methods to establish each of the following prepositions:*
(a) The diagonals of a square are equal in length
(b) The midpoint of the hypotenuse of a right triangle is equidistant from the three vertices;

Problem B.6 * *Use vector methods to determine the equation of the plane through the point $A(1,0,4)$, that is orthogonal to the line through the origin and the point $(3,1,-3)$.*

Problem B.7 * *Use vector methods to prove that: (a) the midpoint of the hypotenuse of a right triangle is equidistant from the three vertices; (b) the diagonals of a rhombus are perpendicular to each other; (c) the base angles of an isosceles triangle are equal to each other.*

3. Conics

A conic is defined by a straight line, called directrix, a focus, at some distance d from the directrix, and the parameter $0 < e < \infty$. The conic is a locus of points with a ratio (e) of distances between the point and directrix and the point and the focus. If we chose the origin in the focus, the equation of a conic is

$$r + e\mathbf{r}\cdot\mathbf{n} = l, \tag{163}$$

where $\mathbf{n}$ is a unit vector perpendicular to the directrix and $l = ed$. If we chose a polar system with the axis OX perpendicular to the directrix, the equation of a conic can be written as

$$r(1 + e\cos\theta) = l. \tag{164}$$

If $e > 1$, we get a hyperbola, if $e = 1$, we get a parabola, if $e < 1$, we get an ellipse.

C. Vector Spaces

Definition C.1 *A group* $(G, *)$ *is a set* G *together with an operation,* $*$*, which to any two elements* x *and* y *of* G *associates element* $z = x * y$ *of* G*. The structure satisfies the following conditions:*

1. *The operation is associative:* $x*(y*z) = (x*y)*z$.
2. *There exists a unique neutral or identity element* n*, such that* $x * n = n * x = x$ *for any* x.
3. *For any element* x *there is a unique element* y *such that* $x * y = y * x$.

If $x * y = y * x$ *for all* x, y *the group is called Abelian or commutative.*

Definition C.2 *A field* F *is a set of elements together with two operations* $+, \times$*.The structure satisfies the following conditions:*

1. $(F, +)$ *is a commutative group.*
2. *The elements of* F *other then* 0 *(neutral element for addition) form a group under multiplication. (Neutral element for multiplication is typically designated as* 1*.)*
3. *Multiplication distributes over addition. That is* $a\times(b+c) = a\times b + a\times c$*, and* $(b+c)\times a = b\times a + c\times a$.

The most important fields are rationals, reals and complexes.

Problem C.1 *For each of the following sets, determine whether they form a group under the operation indicated.*
(a) The integers (mod 10) under addition.
(b) The integers (mod 10) under multiplication.
(c) The integers 1,2,3,4,5,6 under multiplication (mod 7).
(d) The integers 1,2,3,4,5 under multiplication (mod 6).
(e) All matrices of the form

$$\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix}$$

where a *and* b *are integers (mod 5), and* $a \neq 0 \neq b$*, under matrix multiplication.*
(f) All matrices of the form

$$\begin{pmatrix} a & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$$

where a, b, c *are integers, under matrix multiplication.*

Problem C.2 *Show that the set of all numbers of the form* $\alpha + \beta\sqrt[3]{2} + \gamma\sqrt[3]{4}$*, where* α, β, γ *are rational, constitutes a field.*

Definition C.3 *A vector space consists of a field* F*, whose elements are called scalars, an additive Abelian group* V*, whose elements are called vectors, and an operation, called multiplication by scalar, which to any elements* α *of* F *and* X *of* V *associates element* $Y = \alpha \cdot X$ *of* V*, such that the following conditions are satisfied:*

1. *Multiplication by scalars is associative:* $\alpha\cdot(\beta\cdot X) = (\alpha\times\beta)\cdot X$.
2. *Multiplication by scalars is distributive with respect to vector addition:* $\alpha\cdot(X+Y) = \alpha\cdot X + \alpha\cdot Y$.
3. *Multiplication by scalars is distributive with respect to scalar addition:* $(\alpha+\beta)\cdot X = \alpha\cdot X + \beta\cdot X$.
4. $1 \cdot X = X$.

Unless explicitly stated otherwise, the field in this course will be reals or complexes.

Examples of vector spaces are:

1. The set of all n-vectors over some field F, with the operations defined as in matrix algebra. Such space we'll designate F^n. The most important examples are R^n and C^n.
2. The set of geometric vectors, with the operations defined in the beginning of the present Section.
3. The set of all quadratic polynomials $q = \alpha_1 x^2 + \alpha_2 xy + \alpha_3 y^2$, with the operations defined as in school algebra. (The zero element is the polynomial $0x^2 + 0xy + 0y^2$.)
4. The set of all $n \times n$ matrices, with the operations defined as in matrix algebra.
5. The set of all solutions of a given system of homogeneous equations.

Problem C.3 *(a) Consider the set of vectors* (x, y, z) *such that* $x + y + z = 0$ *and* $x + 2y - z = 0$*. Show that these vectors form a vector space.*
(b) Consider the set of vectors (x, y, z) *such that* $x + y + z = 1$ *and* $x + 2y - z = 0$*. Show that these vectors do not form a vector space.*

Problem C.4 *Prove that the set of all real 3-vectors* X *such that* $X^T A = 0$*, where* A *is a fixed real 3-vector, is a vector space over the field of real numbers. Interpret geometrically in* $\mathcal{E}^3$.

D. Linear Combinations and Linear Dependence

Definition D.1 *Let $X_1, X_2, \ldots, X_k$ be vectors in V. The expression $t_1X_1 + t_2X_2 + \ldots + t_kX_k$ is called a linear combination of $X_1, X_2, \ldots, X_k$.*

Definition D.2 *If there exist a linear combination equal to zero, with scalars not all zero, vectors $X_1, X_2, \ldots, X_k$ are called linearly dependent.*

Problem D.1 *Decide whether the vectors $(1,0,0,1)$, $(1,-1,0,0)$, $(0,0,1,0)$,$(1,-1,-1,0)$ are linearly independent.*

Problem D.2 *Show that the pair of vectors $\alpha = (1,3)$ and $\beta = (5,7)$ is linearly independent and find numbers r and s such that $r\alpha + s\beta = (-1,5)$.*

Problem D.3 *For what numbers x will the pair of vectors $(2,3)$ and $(x,5)$ be linearly dependent?*

E. Basis and Dimension

Definition E.1 *We call a vector space V an n-dimensional space if there exists a set of n linearly independent vectors $(V_1, \ldots, V_n)$ that spans the space. That is every vector X of V can be written as*

$$X = x_1V_1 + \ldots + x_nV_n \tag{165}$$

Such set is called basis.

The choice of basis vectors can be done in many ways.

Theorem E.1 *If $V_1, \ldots, V_n$ is a basis for a vector space V, if $U = \alpha_1V_1 + \ldots + \alpha_nV_n$, and if $\alpha_i \neq 0$, then the set $V_1, \ldots, V_{i-1}, U, V_{i+1}, \ldots, V_n$ is also a basis for V.*

Proof E.1 *If*

$$X = x_1V_1 + \ldots + x_nV_n, \tag{166}$$

then

$$X = \left(x_1 - \frac{\alpha_1 x_i}{\alpha_i}\right)V_1 + \ldots + \left(x_{i-1} - \frac{\alpha_{i-1}x_i}{\alpha_i}\right)V_{i-1} + \frac{x_i}{\alpha_i}U + \left(x_{i+1} - \frac{\alpha_{i+1}x_i}{\alpha_i}\right)V_{i+1} + \ldots + \left(x_n - \frac{\alpha_n x_i}{\alpha_i}\right)V_n. \tag{167}$$

Theorem E.2 *If a vector space V has dimension k, then any k linearly independent vectors of V form a basis for V.*

Proof E.2

Any vector X of an n-dimensional vector space can be represented by an n-vector $X = (x_1, \ldots, x_n)^T$. And if we know the algebra of n-vector spaces, we know the algebra of any vector space.

For example, considering geometric vectors on a plane we can chose any two non-collinear vectors as a basis. Hence, the dimensionality of the space is 2. In the space of all n-vectors it is convenient to chose as a basis the set of elementary n-vectors. Such basis is called standard or natural. The number of elements in this basis, and, hence, the dimensionality of the space of all n-vectors is n.

F. Computation of the Dimension of a Vector Space

Very often we define the vector space as the space spanned by a finite set of vectors.

Theorem F.1 *Let we have a set of vectors $V_1, V_2, \ldots, V_k$. The set of all linear combinations these vectors*

$$\alpha_1V_1 + \alpha_2V_2 + \ldots + \alpha_kV_k \tag{168}$$

is a vector space.

Proof F.1

To determine the dimension of the vector space it is convenient to transform the set $V_1, V_2, \ldots, V_k$ into a column echelon form using the elementary column operations:
1. multiplying a vector by a nonzero constant;
2. interchanging two vectors;
3. adding any constant multiple of one vector to another vector.
The operations do not change the space spanned by the vectors. After the set is reduced to a column echelon form, it is obvious, that the non-zero vectors are linearly independent (and can serve as a basis for the space). Their number is the dimensionality of the space.

Let us analyze, for example, the subspace spanned by the columns of the matrix

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & 0 \\ -1 & 4 & -5 & -2 \\ 3 & 3 & 0 & 1 \end{bmatrix}. \tag{169}$$

Sweeping the first (and then the second) row we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ -1 & 6 & -6 & -2 \\ 3 & -3 & 3 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ -1 & 6 & 0 & 0 \\ 3 & -3 & 0 & 0 \end{bmatrix}. \tag{170}$$

So, the dimensionality of the subspace is 2, and we can choose

$$\begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 6 \\ -3 \end{bmatrix} \tag{171}$$

as basis vectors.

Problem F.1 *Find a basis for the space spanned*
(a) by vectors

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -8 \\ 0 \end{bmatrix}.$$

(b) by vectors

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

G. The Inner Product

We may introduce additional structure in a vector space: a scalar function of two vectors X and Y, which is called the inner product and denoted $< X|Y >$ (or (X,Y), or $X \cdot Y$). The function should have the following properties

$$\begin{aligned} &< X|Y > = < Y|X >^* \\ &< X|\lambda Y + \mu Z > \\ &= \lambda < X|Y > + \mu < X|Z > . \end{aligned} \tag{172}$$

Such space is called Hilbert space (or unitary space). We'll designate n-dimensional unitary space U^n.

Problem G.1 *Show, that the space of n-vectors with complex elements, with inner product defined by the equation*

$$< A|B > = \sum_{i=1}^{n} a_i^* b_i$$

is unitary. (This will be our default definition of inner product in C^n.)

If the vector space is over real numbers and the scalar product is real the space is called Euclidean space and designated $\mathcal{E}^n$. In this case $< X|Y > = < Y|X >$. For example, the space of geometric vectors, with inner product defined as scalar multiplication in vector algebra.

Problem G.2 *Show, that the space of n-vectors with real elements, with inner product defined by the equation*

$$< A|B > = \sum_{i=1}^{n} a_i b_i$$

is unitary. (This will be our default definition of inner product in R^n.)

Problem G.3 *Show that the vectors (1,-5,7,2,3) and (2,1,-2,7,1) are orthogonal.*

Problem G.4 *Find*
(a) a linear combination of

$$\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \quad \text{that is orthogonal to} \quad \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix};$$

(b) a linear combination of

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{that is orthogonal to} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Definition G.1 *A norm for for vector X, denoted $\|X\|$ is a real-valued function of vector defined as*

$$\|X\| = \sqrt{< X|X >}. \tag{173}$$

If we divide any vector by it's norm, we get a normalized vector.

For the geometric vectors the norm is the length of the vector.

Problem G.5 *Find the norm of the 4-vector $X = (1 - i, 1 + i, 1, 0)$; then normalize X.*

H. Orthonormal Bases

Let us now introduce an orthonormal basis

$$< V_i|V_j > = \delta_{ij}. \tag{174}$$

In this case we have

$$< X|Y > = \sum_{i=1}^{n} x_i^* y_i. \tag{175}$$

For example, as an orthonormal basis in U^n we may take the natural basis.

Starting from any basis in Hilbert (Euclidean) space $\{V_i\}$ we can get an orthonormal basis $\{U_i\}$, using the Gram-Schmidt orthogonalization procedure:

$$U_1 = \frac{V_1}{\|V_1\|}$$

$$W_j = V_j - \sum_{i=1}^{j-1} (V_j, U_i) V_i \tag{176}$$

$$U_j = \frac{W_j}{\|W_j\|}. \tag{177}$$

For example, we find an orthonormal basis for the vector space spanned by

$$V_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \tag{178}$$

We have first

$$U_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}. \tag{179}$$

Then

$$W_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \left(\left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right] \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}. \tag{180}$$

Hence

$$U_2 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix}. \tag{181}$$

Problem H.1 *Use the Gram-Schmidt process to construct orthonormal bases for the subspaces of $\mathcal{E}^4$ having the following bases:*

$$(a) \quad V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix},$$

$$(b) \quad V_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Problem H.2 *Find a basis for the subspace of all vectors in $\mathcal{E}^4$ that are orthogonal to $(1, 1, 2, 1)$ and $(2, 0, 1, 0)$.*

V. THE RANK OF A MATRIX

A. The Rank of a Matrix

Definition A.1 *The column rank of a matrix is defined to be a dimensionality of its column space.*

Definition A.2 *The row rank of a matrix is defined to be a dimensionality of its row space.*

The row rank of a matrix A and a basis for the row space may be determined by reducing A to a row-echelon form. The column rank of a matrix A and a basis for the column space may be determined by reducing A to a column-echelon form.

Theorem A.1 *The column rank and the row rank of a matrix are the same.*

Proof A.1 *Let us first prove that if the equations*

$$AX = 0 \tag{182}$$

and

$$BX = 0 \tag{183}$$

have the same solution set, then the $n \times n$ matrices A and B have the same column rank. The equations can be presented as

$$\begin{aligned} x_1 A_1 + \ldots + x_n A_n &= 0 \\ x_1 B_1 + \ldots + x_n B_n &= 0. \end{aligned} \tag{184}$$

Let column rank of A is a (we can assume that c first columns are independent), and the column rank of B is b. Suppose that $b < a$. There exist coefficients $d_1, \ldots, d_a$ such that

$$d_1 B_1 + \ldots + d_a B_a = 0 \tag{185}$$

Then $(d_1, \ldots, d_a, 0, \ldots, 0)$ is the solution of Eq. (183) and hence is the solution of Eq. (182). Hence we have

$$d_1 A_1 + \ldots + d_a A_a = 0 \tag{186}$$

in contradiction to the initial assumption. The obvious corollary is that elementary row operations do not change the column rank of a matrix. Now we can prove the theorem. Assume that the column rank of $m \times n$ matrix is c and the row rank is r.

$$B = \begin{bmatrix} A_1 \\ \ldots \\ A_r \end{bmatrix} \tag{187}$$

$$A = \begin{bmatrix} B \\ C \end{bmatrix} = \begin{bmatrix} B \\ TB \end{bmatrix} \tag{188}$$

$$AX = \begin{bmatrix} BX \\ TBX \end{bmatrix} \tag{189}$$

Hence $AX = 0$ if and only if $BX = 0$ and A and B have the same column rank c. But columns of B are r-dimensional vectors, hence

$$c \leq r. \tag{190}$$

Similarly we can prove

$$r \leq c. \tag{191}$$

Hence

$$r = c. \tag{192}$$

Definition A.3 *The common value of the column rank and the row rank of a matrix A is called the rank of a matrix.*

Definition A.4 *If a matrix A has an $r \times r$ submatrix S with $|S| \neq 0$ but no $(r+1) \times (r+1)$ submatrix with non-zero determinant, then r is defined to be a determinant rank of A.*

Theorem A.2 *The determinant rank of a matrix A is equal to its rank.*

Proof A.2 *To prove the theorem it is enough to prove that if A is an $m \times n$ matrix of rank r, then*
(i)Every submatrix C is of rank $\leq r$;
(ii) At least one $r \times r$ submatrix is of rank exactly r.
To prove (i) we can reduce A to C in two stages. The first keeps the number of columns intact and removes only the rows. The row space of the intermediate matrix B is contained in the row space of A, so that rank$(B) \leq$ rank$(A) = r$. At the next stage we remove unwanted columns. We get rank$(C) \leq$ rank$(B) \leq r$.
To prove (ii) suppose that B is formed from r independent rows of A. Then the row space of B is of dimension r; rank$B = r$, and the column space of B must also have dimension r. Then the matrix C formed from r independent columns of B is an $r \times r$ matrix of rank r.

Theorem A.3 *An $n \times n$ matrix A has rank n if and only if A^{-1} exists.*

Proof A.3 *Suppose an $n \times n$ matrix A has rank n. The columns of A then form a basis for the set of all n-vectors, and each vector E_j can be expanded with respect to this basis, that is each equation $AX_j = E_j$ has a unique solution for X_j. Then, since corresponding columns are equal,*

$$[AX_1, AX_2, \ldots, AX_n] = [E_1, E_2, \ldots, E_n], \tag{193}$$

or

$$AX = I. \tag{194}$$

Similarly, expanding with respect to rows we can find

$$Y^T A = I. \tag{195}$$

We have

$$Y^T = Y^T(AX) = IX = X. \tag{196}$$

Hence

$$X = A^{-1}. \tag{197}$$

Suppose, conversely, that A^{-1} exists and put

$$A^{-1} = [C_1, C_2, \ldots, C_n]. \tag{198}$$

We have

$$AC_1 = E_1, \quad AC_2 = E_2, \ldots, AC_n = E_n. \tag{199}$$

Thus each of the n linearly independent vectors $E_1, E_2, \ldots, E_n$ is a linear combination of the columns of A , which are therefore independent also, so that A has rank n.

Problem A.1 *Find the rank of each matrix*

$$(a) \begin{bmatrix} 3 & 2 & 3 & 1 \\ 4 & 3 & 5 & 2 \\ 2 & 1 & 1 & 0 \end{bmatrix} \quad (b) \begin{bmatrix} 0 & 6 & 6 & 1 \\ -8 & 7 & 2 & 3 \\ -2 & 3 & 0 & 1 \\ -3 & 2 & 1 & 1 \end{bmatrix} \quad (c) \begin{bmatrix} 0 & i & -i \\ -i & 0 & i \\ i & -i & 0 \end{bmatrix}$$

Problem A.2 *Under what conditions on x will these matrices:*

$$(a) \begin{bmatrix} x & \sqrt{2} & 0 \\ \sqrt{2} & x & \sqrt{2} \\ 0 & \sqrt{2} & x \end{bmatrix} \quad (b) \begin{bmatrix} 4-x & 2\sqrt{5} & 0 \\ 2\sqrt{5} & 4-x & \sqrt{5} \\ 0 & \sqrt{5} & 4-x \end{bmatrix}$$

fail to have an inverse? (Examine the rank.)

Problem A.3 * *Suppose A is $m \times n$ and B is $n \times m$ matrices, with $n < m$. Prove that their product AB is singular.*

Problem A.4 * *Determine the ranks of the $n \times n$ matrices*

$$(a) \begin{vmatrix} n-1 & 1 & 1 & \ldots & 1 \\ 1 & n-1 & 1 & \ldots & 1 \\ 1 & 1 & n-1 & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & 1 & \ldots & n-1 \end{vmatrix},$$

$$(b) \begin{vmatrix} 1-n & 1 & 1 & \ldots & 1 \\ 1 & 1-n & 1 & \ldots & 1 \\ 1 & 1 & 1-n & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & 1 & \ldots & 1-n \end{vmatrix}.$$

Problem A.5 * *Show that if $AB = 0$, where A and B are of order n and where neither A nor B is the zero matrix, then both A and B are singular.*

Problem A.6 * *Using consideration of rank, show that if $AX = 0$ for all n-vectors X, then $A = 0$.*

B. The Rank of a Matrix and System of Linear Equations

Theorem B.1 *The complete solution of a system of homogeneous linear equations in n unknowns is a linear combination of $(n-r)$ linear independent particular solutions, where r is the rank of the matrix A.*

Proof B.1

Theorem B.2 *The system of nonhomogeneous equations is consistent if and only if the rank of A equals the rank of* $[\mathbf{A}|\mathbf{B}]$.

Proof B.2

Theorem B.3 *If the system of m linear equations in n unknowns represented by $AX = B$ is consistent and of rank r, then any r linearly independent equations of this system form an equivalent system.*

Proof B.3

Theorem B.4 *If $AX = B$ is consistent and of rank r, we can solve for $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ in terms of the remaining variables if and only if the $m \times r$ submatrix of A which contains the coefficients of $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ has rank r.*

Proof B.4

Problem B.1 *Show, that one can solve system*

$$\begin{aligned} x_1 - 2x_2 + 3x_3 - x_4 &= 1 \\ 2x_1 - 4x_2 + 5x_3 - 2x_4 &= 0 \end{aligned}$$

for x_1 and x_3, x_2 and x_3, or x_3 and x_4, but not for x_1 and x_2, x_1 and x_4, or x_2 and x_4.

Problem B.2 *For which set of variables can one solve the following system?*

$$\begin{aligned} x_1 + 2x_2 - x_3 - 2x_4 &= 1 \\ 2x_1 - x_2 + 3x_3 + x_4 &= -3 \\ 3x_1 + 2x_2 + 2x_3 - 2x_4 &= -2. \end{aligned}$$

Compute the basic solutions.

VI. LINEAR TRANSFORMATIONS

A. Linear Transformations

Definition A.1 *A transformation, or mapping, or function is a one-way correspondence which associates each element of one system with a unique element of another – or possibly the same – system.*

Definition A.2 *A mapping $\mathcal{A}$ of a vector space V into a vector space W is called a linear transformation of V into W if, for any vectors $\mathcal{X}_1, \mathcal{X}_2 \in V$ and an arbitrary number r*

$$\begin{aligned} &(1)\ \mathcal{A}(\mathcal{X}_1 + \mathcal{X}_2) = \mathcal{A}\mathcal{X}_1 + \mathcal{A}\mathcal{X}_2 \\ &(2)\ \mathcal{A}(r\mathcal{X}_1) = r\mathcal{A}\mathcal{X}_1. \end{aligned} \tag{200}$$

For example, rotation of the plane through an angle θ maps any geometric vector into another geometric vectors. It can be easily proved that the mapping is linear.

B. Linear Transformations and Matrices

Consider an $n-$dimensional space V and $m-$dimensional space W. Let us fix a basis $(\mathcal{V}_1, \ldots, \mathcal{V}_n)$ in V and $(\mathcal{W}_1, \ldots, \mathcal{W}_m)$ in W. There is a one-to-one correspondence between a linear transformation $\mathcal{A}$ and an $m \times n$ matrix A_{ij}, defined by the equation

$$\mathcal{A}\mathcal{V}_i = A_{ji}\mathcal{W}_j. \tag{201}$$

We have

$$\mathcal{Y} = \mathcal{A}\mathcal{X} \Longleftrightarrow Y = AX. \tag{202}$$

For example, rotation of the plane through an angle θ, corresponds to the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}; \tag{203}$$

stretching of the plain in one direction by a factor of 2, and compressing it in a perpendicular direction by a factor of 3, corresponds to the transformation of plane geometric vectors, described by a matrix

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1/3 \end{bmatrix}, \tag{204}$$

if $\mathcal{V}_1$ chosen along the first direction, and $\mathcal{V}_2$ – along the second.

Especially important for us the transformation of a vector space V into itself (such transformation is called a linear operator on V), and the transformation of V into the space $\mathbf{R}$ of real numbers (such transformation is called a linear functional).

Problem B.1 *Describe the effect of each of the transformations on geometric vectors: (a) $T(x_1, x_2) = (x_2, x_1)$; (b) $T(x_1, x_2) = (x_1, 3x_2)$; (c) $T(x_1, x_2) = (x_2 - x_1, x_1 - x_2)$.*

Problem B.2 *Find the transformation which carries the vectors $\{1, 1\}$ and $\{3, -2\}$ into the vectors $\{2, 1\}$ and $\{1, 2\}$ respectively.*

Problem B.3 *If T_1 and T_2 are linear operators defined by $T_1(x_1,x_2,x_3) = (2x_1 - x_2, x_1, x_2 - x_3)$ and $T_2(x_1,x_2,x_3) = (x_1 - x_3, x_2 + x_3, 0)$, determine (a) $T_1 + T_2$; (b) T_1T_2; (c) T_2T_1; (d) $2T_1$; (e) $2T_1 + 3T_2$.*

Problem B.4 *Find a matrix representation of the operator that rotates each vector in $\mathcal{E}^2$ through an angle (a) 2 rad; (b) $30°$; (c) $60°$.*

C. Change of an Operator Matrix under Change of Basis

Let we have a basis $(\mathcal{V}_1,\ldots,\mathcal{V}_n)$. Let us introduce a new basis $(\mathcal{V}'_1,\ldots,\mathcal{V}'_n)$, which is related to the old basis, by the transformation

$$\mathcal{V}'_j = \sum_{i=1}^{n} S_{ij}\mathcal{V}_i. \tag{205}$$

For an arbitrary vector $\mathcal{X}$ it follows, that

$$\mathcal{X} = \sum_{i=1}^{n} x_i\mathcal{V}_i = \sum_{j=1}^{n} x'_j\mathcal{V}'_j = \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} S_{ij}\mathcal{V}_i. \tag{206}$$

Hence

$$x_i = \sum_{j=1}^{n} S_{ij}x'_j, \tag{207}$$

which we can present in matrix form

$$X = SX', \tag{208}$$

or equivalently

$$X' = S^{-1}X. \tag{209}$$

Consider linear operator, which can be presented in each of two bases

$$\begin{aligned} Y &= AX \\ Y' &= AX'. \end{aligned} \tag{210}$$

After simple algebra from (208) we get

$$Y' = S^{-1}ASX'. \tag{211}$$

Hence

$$A' = S^{-1}AS. \tag{212}$$

Given a square matrix A, representing the linear operator A, we may also consider the matrix $A' = S^{-1}AS$ for any nonsingular matrix S as representing the same linear operator but in a new basis.

Problem C.1 *Obtain the matrix of the coordinate transformation in the space of n-vectors which replaces the natural basis by the basis*

$$B_k = \sum_{j=1}^{k} E_j, \qquad k = 1,2,\ldots,n.$$

D. Unitary Operators

Definition D.1 *Matrix A is called unitary if and only if*

$$AA^\dagger = I. \tag{213}$$

Definition D.2 *A real unitary matrix is called orthogonal.*

Problem D.1 *Find, whether matrix*

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{214}$$

is orthogonal.

Definition D.3 *An operator in unitary space which is presented by a unitary matrix (in an orthonormal basis) is called a unitary operator.*

Definition D.4 *An operator in Euclidean space which is presented by an orthogonal matrix (in an orthonormal basis) is called an orthogonal operator.*

Theorem D.1 *A linear operator in unitary space leaves the norms of all vectors invariant if and only if it is a unitary operator.*

Proof D.1 *For any vector*

$$(AX)^\dagger(AX) = X^\dagger A^\dagger AX = X^\dagger(A^\dagger A)X. \tag{215}$$

If A is unitary then

$$X^\dagger(A^\dagger A)X = X^\dagger X. \tag{216}$$

Hence the norm is invariant. Suppose now, that Eq. (216) is true for any vector. Let us take $X = E_j$. We get

$$(A^\dagger A)_{jj} = 1. \tag{217}$$

Theorem D.2 *A linear operator in Euclidean space leaves the norms of all vectors invariant if and only if it is an orthogonal operator.*

Proof D.2

Problem D.2 *Determine necessary and sufficient conditions for the matrix*

$$\begin{bmatrix} x+y & y-x \\ x-y & y+x \end{bmatrix}$$

to be orthogonal.

Problem D.3 * *Prove that if S is a real antisymmetric matrix then*

$$A = (I-S)(I+S)^{-1}$$

is orthogonal and express the matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

in this form.

Problem D.4 *A and B are real non-zero 3×3 matrices and satisfy the equation*

$$(AB)^T + B^{-1}A = 0.$$

(a) Prove that if B is orthogonal then A is antisymmetric.
(b) Without assuming that B is orthogonal, prove that A is singular.

VII. EIGENVALUES AND EIGENVECTORS

A. Definition of the Eigenvalue Problem

Definition A.1 *The eigenvector (or characteristic vector) of matrix A is a vector X which satisfies equation*

$$AX = \lambda X. \tag{218}$$

The scalar λ is called an eigenvalue (or characteristic value).

If we rewrite Eq. (218) in the form

$$(A-\lambda I)X = 0, \tag{219}$$

We see that λ is a solution of the characteristic equation

$$\det(A-\lambda I) = 0. \tag{220}$$

The polynomial

$$\phi(\lambda) = \det(A-\lambda I) \tag{221}$$

is called the characteristic polynomial.
Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}. \tag{222}$$

The eigenvalues are the solution of the equation

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) = 0. \tag{223}$$

The eigenvector corresponding to $\lambda = 1$ is

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0. \tag{224}$$

In a scalar form the system reduces to a single equation

$$x_1 + x_2 = 0, \tag{225}$$

which has the complete solution

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{226}$$

The eigenvector corresponding to $\lambda = 3$ is

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0. \tag{227}$$

In a scalar form the system reduces to a single equation

$$x_1 - x_2 = 0, \tag{228}$$

which has the complete solution

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{229}$$

Theorem A.1 *Let $\lambda_1, \lambda_2, \ldots \lambda_k$ be distinct eigenvalues of a matrix A, and let $X_1, X_2, \ldots X_k$ be eigenvectors. Then $X_1, X_2, \ldots X_k$ are linearly independent*

Proof A.1 *Suppose it is not so, that is*

$$\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k = 0 \tag{230}$$

Let us act on this equation by the operator $(A-\lambda_2)(A-\lambda_3)\ldots(A-\lambda_k)$. We get

$$\alpha_1(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\ldots(\lambda_1-\lambda_k) = 0, \tag{231}$$

that is $\alpha_1 = 0$. We thus can show that any α is equal to zero, that is our assumption is wrong.

Problem A.1 *Find the eigenvalues and eigenvectors of the matrices*

$$(a)\ \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}; \quad (b)\ \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}; \quad (c)\ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Problem A.2 *Prove that the matrices A and A^T have the same eigenvalues.*

Problem A.3 *If $A = [a_{ij}]$ is a triangular matrix, prove that the eigenvalues are the numbers along the main diagonal.*

Problem A.4 * *Prove that a matrix is singular if and only if it has a zero eigenvalue.*

B. Similar Matrices

Definition B.1 *A matrix C is similar to A if there exists a matrix B such, that*

$$C = B^{-1}AB. \tag{232}$$

If matrix B can be chosen unitary, the matrices A and C are called unitary similar.

Theorem B.1 *Let A and C be similar matrices. Then A and C have identical characteristic polynomials and hence have identical eigenvalues. If X is an eigenvector of A then $B^{-1}X$ is an eigenvector of C corresponding to the same eigenvalue.*

Proof B.1

Problem B.1 *Determine which of the following pairs are similar matrices*

$$\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 & 7 \\ 7 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

C. Diagonalization of a Matrix

Theorem C.1 *There exists nonsingular matrix S such, that*

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_N \end{bmatrix} \tag{233}$$

if and only the $n \times n$ matrix A possess n linearly independent characteristic vectors.

Proof C.1 *Consider matrix*

$$S = [V_1, V_2, \dots, V_n], \tag{234}$$

where V_i is the solution of the equation

$$AV_i = \lambda_i V_i. \tag{235}$$

Then

$$AS = S \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_N \end{bmatrix}. \tag{236}$$

Because the vectors $\{V_1, V_2, \dots, V_n\}$ are linearly independent, S^{-1} exists, and we obtain Eq. (233).

Problem C.1 *Prove that, if B is nonsingular and commutes with A, and if U diagonalizes A, then BU also diagonalizes A.*

There exist matrices which cannot be reduced to a diagonal form by means of a nonsingular matrix. Equivalently, there exist matrices of dimension N which do not possess N linearly independent characteristic vectors. Consider two matrices

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{237}$$

Both have a single eigenvalue $\lambda = 1$, but the first has two linearly independent eigenvectors

$$X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{238}$$

and the second has only one eigenvector X_1, and hence can not be diagonalized.

D. The Eigenvalues of a Hermitian Matrix

Definition D.1 *A transposed conjugate (or tranjugate, or adjoint) of a matrix A is designated $A^\dagger$ and is defined*

$$[A^\dagger]_{ij} = [A]^*_{ji} \tag{239}$$

Problem D.1 *Show that*

$$(AB)^\dagger = B^\dagger A^\dagger.$$

Definition D.2 *A matrix is called Hermitian, if*

$$A^\dagger = A \Longleftrightarrow a_{ij} = a^*_{ji}. \tag{240}$$

Problem D.2 *Check up whether the matrix*

$$\begin{bmatrix} 1 & 2+3i \\ 2-3i & 4 \end{bmatrix}$$

is Hermitian.

Problem D.3 *Prove that if A and B are Hermitian and $AB = BA$, then AB is Hermitian. What does this imply about powers of Hermitian matrices?*

Theorem D.1 *The eigenvalues of a Hermitian matrix are all real*

Proof D.1 *Let λ be an eigenvalue and Y the corresponding eigenvector*

$$AY = \lambda Y. \tag{241}$$

Then

$$Y^\dagger AY = \lambda Y^\dagger Y = Y^\dagger A^\dagger Y = \left(Y^\dagger AY\right)^\dagger = \lambda^* Y^\dagger Y. \tag{242}$$

Hence λ is real.

Theorem D.2 *If X and Y are eigenvectors associated with distinct eigenvalues of a Hermitian matrix, then X and Y are orthogonal.*

Proof D.2 *Suppose, in fact, that*

$$AX = \lambda X \ \text{ and } \ AY = \mu Y, \ \ \lambda \neq \mu \tag{243}$$

$$Y^\dagger AX = \lambda Y^\dagger X \tag{244}$$

$$X^\dagger AY = \mu X^\dagger Y \to Y^\dagger AX = \mu Y^\dagger X \tag{245}$$

$$\lambda Y^\dagger X = \mu Y^\dagger X \tag{246}$$

Since $\lambda \neq \mu$ we get $Y^\dagger X = 0$.

Problem D.4 *Find the eigenvalues and eigenvectors of the matrices*

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}, \ \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \ \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}, \ \begin{bmatrix} 3 & -2 \\ -2 & 0 \end{bmatrix}.$$

Problem D.5 *Determine the characteristic roots and vectors of the Hermitian matrix*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \omega^2 \\ 0 & \omega & 0 \end{bmatrix},$$

where ω is a complex cube root of unity: $\omega = e^{2\pi i/3}$.

Problem D.6 *For the matrix*

$$A = \begin{bmatrix} 1 & \alpha & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where α and β are non-zero complex numbers, find the eigenvalues and eigenvectors. Find the respective conditions for (a) the eigenvalues to be real, and (b) the eigenvectors to be orthogonal. Show that the conditions are jointly satisfied only if A is Hermitian.

Problem D.7 *(a) Show that if A is Hermitian and U is unitary then $U^{-1}AU$ is Hermitian.*

(b) Show that if A is anti-Hermitian then iA is Hermitian.

(c) Prove that the product of two Hermitian Matrices A and B is Hermitian if and only if A and B commute.

Problem D.8 ** Prove that the characteristic roots of $A^\dagger A$ are all non-negative.*

E. The Diagonalization of a Hermitian Matrix

Theorem E.1 *With every Hermitian matrix A we can associate an orthonormal set of n eigenvectors.*

Proof E.1

Theorem E.2 *If $U_1, U_2, \ldots, U_n$ is an orthonormal system of characteristic vectors associated respectively with the characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of a Hermitian matrix A,then $U = [U_1, U_2, \ldots, U_n]$ is the unitary matrix, and*

$$U^\dagger AU = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_N \end{bmatrix}. \tag{247}$$

Proof E.2 *We have*

$$AU = A[U_1, U_2, \ldots, U_n] = [\lambda_1 U_1, \lambda_2 U_2, \ldots, \lambda_n U_n]. \tag{248}$$

Hence

$$U^\dagger AU = \begin{bmatrix} U_1^\dagger \\ U_2^\dagger \\ \ldots \\ U_N^\dagger \end{bmatrix} [\lambda_1 U_1, \lambda_2 U_2, \ldots, \lambda_n U_n]$$

$$= \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_N \end{bmatrix}. \tag{249}$$

Theorem E.3 *If A is a real symmetric matrix there exists an orthogonal matrix U such that $U^{-1}AU$ is a diagonal matrix whose diagonal elements are the eigenvalues of A.*

Proof E.3

Problem E.1 *Find the orthogonal matrices which diagonalize the matrices*

$$(a) \ \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \quad (b) \ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (c) \ \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$(c) \ \begin{bmatrix} 0 & \sqrt{2} & -1 \\ \sqrt{2} & 1 & -\sqrt{2} \\ -1 & -\sqrt{2} & 0 \end{bmatrix}, \quad (d) \ \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}.$$

F. Normal Matrices

Definition F.1 *A matrix A is said to be normal if and only if*

$$A^\dagger A = AA^\dagger \tag{250}$$

Obviously, both Hermitian and unitary matrices are normal.

Problem F.1 *Determine which of the following matrices are normal:*

$$\begin{bmatrix} 1 & 4 & 3 \\ 4 & 4 & -5 \\ 3 & -5 & 0 \end{bmatrix}; \quad \begin{bmatrix} 7 & 3+5i \\ 3-5i & 4 \end{bmatrix}; \quad \begin{bmatrix} 2 & 6 & -3 \\ 3 & 2 & 6 \\ -6 & 3 & 2 \end{bmatrix}.$$

Theorem F.1 *A matrix A over the complex field can be diagonalized by a unitary transformation if and only if A is normal.*

Proof F.1 *Let the matrix A can be diagonalized by a unitary transformation:*

$$A = S \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_N \end{bmatrix} S^{-1}. \tag{251}$$

Then

$$A^\dagger = S \begin{bmatrix} \lambda_1^* & 0 & \dots & 0 \\ 0 & \lambda_2^* & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_N^* \end{bmatrix} S^{-1} \tag{252}$$

and

$$A^\dagger A = S \begin{bmatrix} |\lambda_1|^2 & 0 & \dots & 0 \\ 0 & |\lambda_2|^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & |\lambda_N|^2 \end{bmatrix} S^{-1} = AA^\dagger.$$

Problem F.2 *Find a unitary transformation U such that*

$$U^\dagger \begin{bmatrix} a & -b \\ b & a \end{bmatrix} U = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

where a and b are real numbers.

Problem F.3 *Prove that all eigenvalues of a unitary matrix have modulus 1.*

G. The Cayley-Hamilton Theorem

Theorem G.1 *Every square matrix satisfies its own characteristic equation. That is, if*

$$det(A - \lambda I) = \phi(\lambda) \tag{253}$$

then

$$\phi(A) = 0. \tag{254}$$

Proof G.1 *To prove the theorem consider the matrix inverse to $A - \lambda I$, for λ not a characteristic value.*

$$(A - \lambda I)^{-1} = B(\lambda)/\phi(\lambda). \tag{255}$$

Where $B(\lambda)$ is a matrix adjoint to $A - \lambda I$, and hence a polynomial of the degree $n-1$ with respect to λ.

$$B(\lambda) = \lambda^{N-1}B_{N-1} + \lambda^{N-2}B_{N-2} + \ldots + B_0. \tag{256}$$

We see that the polynomial $\phi(\lambda)I$ is divisible without remainder on $A - \lambda I$

$$\phi(\lambda)I = (A - \lambda I)B(\lambda). \tag{257}$$

Hence Eq. (254).

Problem G.1 *If*

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

check that $T^2 - (a+d)T + (ad - bc)I = 0$.

Problem G.2 *Verify the Cayley-Hamilton theorem for*

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}.$$

H. Functions of Matrices

Definition H.1 *Let $n \times n$ matrix A has n distinct eigenvalues, and function $f(\lambda)$ is defined on the spectrum of matrix A. Then*

$$f(A) = g(A),$$

where $g(\lambda)$ is an arbitrary polynomial that assumes on the spectrum of A the same values as does $f(\lambda)$.

Theorem H.1 *Among all the polynomials with complex coefficients that assume on the spectrum of A the same values as does $f(\lambda)$ there is only one polynomial that is of degree less than n.*

Proof H.1 *Let us write $\phi(A) = 0$ in the form*

$$A^n = -a_0 I - a_1 A - \cdots - a_{n-1}A^{n-1}. \tag{258}$$

Hence any power of matrix A^k $(k \geq n)$ can be expresses as a polynomial of degree less than n in A.

To find explicitly the coefficients b in the expansion

$$f(A) = b_{n-1}A^{n-1} + b_{n-2}A^{n-2} + \ldots + b_1 A + b_0 I \quad (259)$$

we should solve system of linear equations:

$$f(\lambda_i) = b_{n-1}\lambda_i^{n-1} + b_{n-2}\lambda_i^{n-2} + \ldots + b_1\lambda_i + b_0. \quad (260)$$

Problem H.1 *Find $e^{\epsilon\sigma_z}$, $\cos\sigma_z$, and σ_z^{31}.*

Theorem H.2 *If the function $f(z)$ can be expanded in a power series*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \Longrightarrow f(A) = \sum_{n=0}^{\infty} a_n A^n. \quad (261)$$

Proof H.2

Problem H.2 *Calculate $e^{\epsilon\sigma_z}$ directly summing up the series for the exponential*

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n.$$

Hint: Use the obvious relations $\sigma_z^{2k+1} = \sigma_z$ and $\sigma_z^{2k} = I_2$.

Theorem H.3 *If the matrix A is diagonalized*

$$A = S \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_N \end{bmatrix} S^{-1}, \quad (262)$$

then

$$f(A) = S \begin{bmatrix} f(\lambda_1) & 0 & \ldots & 0 \\ 0 & f(\lambda_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & f(\lambda_N) \end{bmatrix} S^{-1}. \quad (263)$$

Proof H.3

Problem H.3 ** Calculate $e^{\epsilon\sigma_x}$ using all three methods. Compare the results.*

VIII. QUADRATIC FORMS

A. Quadratic Forms and Calculus

A homogeneous polynomial q of the type

$$q = \sum_{i,j=1}^{n} a_{ij} x_i x_j \quad (264)$$

is called a quadratic form.

Consider a function of two variables $f(x_1, x_2)$. What are the necessary and sufficient conditions for the fact that the function reaches maximum at some point, say $(0,0)$? To answer this questions we can expand the function about $(0,0)$

$$f(x_1, x_2) = f(0,0) + \left(\frac{\partial f}{\partial x_1}\right)_0 x_1 + \left(\frac{\partial f}{\partial x_2}\right)_0 x_2. \quad (265)$$

It is necessary that

$$\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0. \quad (266)$$

Definition A.1 *The point where all the first derivatives are equal to zero is called a critical point.*

To find the sufficient condition we should expand the function further. The higher terms are

$$\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}\right)_0 x_1^2 + \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)_0 x_1 x_2 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_2^2}\right)_0 x_2^2. \quad (267)$$

We see that the answer is reduced to the analysis of the quadratic form

$$f = a x_1^2 + 2b x_1 x_2 + c x_2^2. \quad (268)$$

We can transform it to the form

$$f = a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(c - \frac{b^2}{a}\right) x_2^2. \quad (269)$$

The condition for the minimum is

$$\begin{aligned} a &> 0 \\ c - b^2/a &> 0. \end{aligned} \quad (270)$$

The condition for the maximum is opposite inequality signs. If the two coefficients have different signs we have a saddle point.

Problem A.1 *Consider the function*

$$f = 2x^2 + 4xy + y^2. \quad (271)$$

What is the nature of the critical point at the origin?

Problem A.2 *Decide between a minimum, maximum, or saddle point for the functions*

(a) $F = -1 + 4(e^x - x) - 5x\sin y + 6y^2$ at the point $x = y = 0$;

(b) $F = (x^2 - 2x)\cos y$ with stationary point at $x = 1$, $y = \pi$.

B. Reduction of a Quadratic Form to a Sum of Squares

Theorem B.1 *A real quadratic form q can be represented in an infinite number of ways in the form*

$$q = \sum_{i=1}^{r} a_i X_i^2,$$

where

$$X_i = \sum_{i=1}^{n} \alpha_{ki} x_k,$$

and the number of positive and negative squares for a given form q are independent of the choice of representation. (Sylvester's Law of Inertia for Real Quadratic Forms.)

Proof B.1

1. The method of Lagrange of Reducing of a quadratic form to a sum of squares

The method involves just a repeated completing of the square. Let a quadratic form

$$q = \sum_{i,j=1}^{n} a_{ij} x_i x_j \tag{272}$$

is given.

We consider two cases:

1) for some g the diagonal coefficient a_{kk} is not equal to zero. Then we set

$$q = \frac{1}{a_{gg}} \left(\sum_{k=1}^{n} a_{gk} x_k \right)^2 + q' \tag{273}$$

and see that the quadratic form q' does not contain x_g.

2) $a_{gg} = 0$ and $a_{hh} = 0$, but $a_{gh} \neq 0$. Then we set

$$q = \frac{1}{2a_{gh}} \left[\sum_{i=1}^{n} (a_{gk} + a_{hk}) x_k \right]^2 - \frac{1}{2a_{gh}} \left[\sum_{i=1}^{n} (a_{gk} - a_{hk}) x_k \right]^2 + q' \tag{274}$$

and see that the quadratic form q' does not contain neither x_g nor x_h.

Consider, for example, the form

$$q = 2x_1^2 + x_1x_2 - 3x_1x_3 + 2x_2x_3 - x_3^2. \tag{275}$$

Simple algebra gives

$$q = \frac{1}{2}\left[2x_1 + \frac{1}{2}x_2 - \frac{3}{2}x_3\right]^2 - \frac{1}{8}x_2^2 + \frac{11}{4}x_2x_3 - \frac{17}{8}x_3^2$$
$$= \frac{1}{2}\left[2x_1 + \frac{1}{2}x_2 - \frac{3}{2}x_3\right]^2 - 8\left[-\frac{1}{8}x_2 + \frac{11}{8}x_3\right]^2 + 13x_3^2. \tag{276}$$

C. Quadratic Forms and Matrices

Quadratic form can be presented in a matrix form

$$q = \sum_{i,j=1}^{n} a_{ij} x_i x_j = X^T A X. \tag{277}$$

Problem C.1 *Check up the equation*

$$x_1^2 + 4x_1x_2 + 4x_2^2 = X^T \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} X.$$

Problem C.2 *Find the matrix of each of the bilinear forms:*

(a) $2x_1y_1 - 3x_1y_3 + 2x_2y_2 - 5x_2y_3 + 4x_3y_1 + x_3y_3$,
(b) $4x_1y_1 + 2x_2y_2 + x_3y_3$,
(c) $x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4$.

A nonsingular linear transformation $X = B\tilde{X}$ maps a quadratic form into another one

$$X^T A X \Longrightarrow \tilde{X}^T (B^T A B) \tilde{X}. \tag{278}$$

Definition C.1 *Two matrices, A_1 and $A_2 = B^T A B$, where B is any nonsingular matrix, are called congruent.*

D. Reduction of a Quadratic Form to Principal axes

For every real symmetric matrix A, there exists an orthogonal matrix U such that

$$U^T A U = D[\lambda_1, \lambda_2, \ldots, \lambda_n], \tag{279}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of A. As a consequence, the transformation $X = U\tilde{X}$ applied to the quadratic form $X^T A X$ gives

$$X^T A X = \lambda_1 \tilde{x}_1^2 + \lambda_2 \tilde{x}_2^2 + \ldots + \lambda_n \tilde{x}_n^2. \tag{280}$$

Problem D.1 *Find an orthogonal matrix P such that $P^T A P$ is in diagonal form, where $A =$*

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}; \begin{bmatrix} -2 & 2\sqrt{3} \\ 2\sqrt{3} & 2 \end{bmatrix}; \begin{bmatrix} 1 & 4\sqrt{5} \\ 4\sqrt{5} & -1 \end{bmatrix}.$$

Check your results.

Problem D.2 *Diagonalize each quadratic form by means of an orthogonal transformation*

(a) $q(X) = 2x_1x_2 + 2x_2x_3$,
(b) $q(X) = 2x_1x_2 + 2x_2x_3 + 2x_3x_1$,

Consider the graph of the quadratic equation

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{21}x_2^2 = b \tag{281}$$

We can rewrite it in the form

$$X^T AX = b. \tag{282}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}. \tag{283}$$

There exists an orthogonal matrix U ($U^T = U^{-1}$), which diagonalizes matrix A

$$U^T AU = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}. \tag{284}$$

Hence, if we put

$$X = UY, \tag{285}$$

we obtain Eq. (282) in the form

$$Y^T \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} Y = b, \tag{286}$$

or

$$\frac{y_1^2}{b/\lambda_1} + \frac{y_2^2}{b/\lambda_2} = 1. \tag{287}$$

Depending upon the signs of b/λ_1 and b/λ_2 we obtain ellipse, hyperbola or an empty set. The new axes y_1 and y_2 are called the principal axes of the conic. Their direction relative to the old axes is given by Eq. (285), which can be presented in the form

$$\begin{aligned} \mathbf{e}_1' &= U_{11}\mathbf{e}_1 + U_{21}\mathbf{e}_2 \\ \mathbf{e}_2' &= U_{12}\mathbf{e}_1 + U_{22}\mathbf{e}_2. \end{aligned} \tag{288}$$

Problem D.3 *The equation of a conic section is*

$$Q \equiv 8x_1^2 + 8x_2^2 - 6x_1x_2 = 110.$$

Determine the type of conic section this represents, the orientation of its principal axes, and relevant lengths in the directions of these axes.

Problem D.4 *Identify each of the following conics and sketch a figure*
(a) $x_1^2 + 6x_1x_2 + x_2^2 = 4$,
(b) $x_1^2 + 2x_1x_2 + 4x_2^2 = 6$,
(c) $4x_1^2 - x_1x_2 + 4x_2^2 = 4$.

Similarly we can consider an arbitrary quadratic surface in $\mathcal{E}^3$ and obtain equation

$$\frac{y_1^2}{b/\lambda_1} + \frac{y_2^2}{b/\lambda_2} + \frac{y_2^3}{b/\lambda_3} = 1 \tag{289}$$

which can give ellipsoid (if all three denominators are positive), hyperboloid of one sheet (if one of the denominators is negative), hyperboloid of two sheets (if two of the denominators are negative), or empty set (if all three denominators are negative).

Problem D.5 *Show that the quadratic surface*

$$5x^2 + 11y^2 + 5z^2 - 10yz + 2xz - 10xy = 4$$

is an ellipsoid with semi-axes of 2, 1 and 0.5. Find the direction of the longest axis.

E. Definite Quadratic Forms

Definition E.1 *A real quadratic form q defined by $q(X) = X^T AX$, is positive definite if $q(X) > 0$ for all real X except $X = 0$.*

Definition E.2 *A real quadratic form q defined by $q(X) = X^T AX$, is positive semidefinite (or nonnegative) if $q(X) \geq 0$ for all real X.*

Definition E.3 *A real quadratic form q defined by $q(X) = X^T AX$, is negative definite if $q(X) < 0$ for all real X except $X = 0$.*

Definition E.4 *A real quadratic form q defined by $q(X) = X^T AX$, is negative semidefinite (or nonpositive) if $q(X) \leq 0$ for all real X.*

Theorem E.1 *A quadratic form $X^T AX$ is positive definite if and only if the characteristic roots of A are all positive.*

Proof E.1 *Assume that the form is positive definite. Let λ_1 (necessary real) be any characteristic root of A.*

$$AU_1 = \lambda_1 U_1. \tag{290}$$

Then since the form is positive definite

$$0 < U_1^T AU_1 = \lambda_1 U_1 = \lambda_1. \tag{291}$$

Now assume that all the characteristic roots are positive and $X = \sum \alpha_i U_i$. Hence

$$X^T AX = \sum_i \alpha_i^2 \lambda_i. \tag{292}$$

Definition E.5 *Consider a $n \times n$ matrix A. Every principal minor of order r is the determinant of a submatrix obtainable by deleting symmetrically a set of rows and columns of A. The minor*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{bmatrix} \tag{293}$$

is called the leading principal minor of order r.

Theorem E.2 *A real quadratic form is positive definite if and only if the leading principal minors of the matrix of the form are all positive.*

Proof E.2 *If the form is positive definite, then $detA = \lambda_1 \dots \lambda_n > 0$. Every principal minor of order r corresponds to putting $n-r$ of the x_i equal to zero and looking at the quadratic form of the remaining variables only. Since the original form is positive definite so is the reduced form.*

Theorem E.3 *A real quadratic form is positive definite if and only if it can be reduced to upper triangular form using only elementary row operations E3 and the diagonal elements of the resulting matrix (the pivots) are all positive.*

Proof E.3

Useful necessary (not sufficient) conditions for positive definiteness are:

1. every diagonal element $a_{ii} > 0$;
2. the element of matrix having the greatest absolute value must be on the diagonal;
3. for any i and $j$the inequality $a_{ii}a_{jj} > |a_{ij}|^2$ holds.

Problem E.1 *Examine for definiteness:*

$$(a) \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & 4 \\ -1 & 4 & 6 \end{bmatrix}; (b) \begin{bmatrix} 4 & 2 & -2 \\ 2 & 4 & 2 \\ -2 & 2 & 4 \end{bmatrix}; (c) \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 0 & 3 & -3 \end{bmatrix}.$$

Problem E.2 *Show that for every real matrix A, A^TA is positive semidefinite or positive definite. When is it positive definite?*

Problem E.3 *** Under what conditions on α will the following tridiagonal matrix be positive definite?*

$$\begin{bmatrix} \alpha & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & \alpha & -1 & 0 & \dots & \dots & 0 \\ 0 & -1 & \alpha & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & -1 & \alpha & -1 \\ 0 & \dots & \dots & \dots & \dots & -1 & \alpha \end{bmatrix}_{n\times n}$$

is positive definite.

F. Simultaneous Reduction of two Quadratic Forms to the Sums of Squares and Generalized Eigenvalue Problem

Theorem F.1 *If A and B are real symmetric matrix of order n and B is positive definite, then there exists a real nonsingular matrix V such that V^TAV is diagonal and V^TBV is an identity matrix.*

Proof F.1 *There exists an orthogonal matrix U such that*

$$U^TBU = D[\mu_1, \mu_2, \dots, \mu_n], \tag{294}$$

where $\mu_1, \dots, \mu_n$ are all positive. If we put

$$R = D[\mu_1^{-1/2}, \mu_2^{-1/2}, \dots, \mu_n^{-1/2}], \tag{295}$$

and $S = UR$, we have

$$S^TBS = I. \tag{296}$$

Since S^TAS is a real symmetric matrix, there exists an orthogonal matrix Q such, that

$$Q^TS^TASQ = (SQ)^TASQ = D[\lambda_1, \lambda_2, \dots, \lambda_n], \tag{297}$$

that is

$$V^TAV = D[\lambda_1, \lambda_2, \dots, \lambda_n], \tag{298}$$

where $V = SQ$. Also

$$V^TBV = I. \tag{299}$$

Definition F.1 *The generalized eigenvector of matrix A is a vector X which satisfies equation*

$$AX = \lambda BX. \tag{300}$$

The scalar λ is called a generalized eigenvalue.

Theorem F.2 *1. The eigenvalues are the solutions of the equation*

$$\det[A - \lambda B] = 0. \tag{301}$$

2. If A and B are symmetric and B is positive definite, all eigenvalues are real.

Proof F.2 *We may replace Eq. (301) by*

$$det[S^TAS - \lambda I] = 0. \tag{302}$$

S^TAS is a real symmetric matrix, hence all its eigenvalues are real

Theorem F.3 *If A and B are real symmetric matrix of order n and A has n distinct B-generalized eigenvalues, the matrix built from generalized eigenvectors diagonalizes both A and B.*

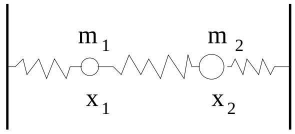
Proof F.3 *Consider two eigenvalues $\lambda_1 \neq \lambda_2$. We get*

$$\begin{cases} AV_1 = \lambda_1 BV_1 \\ AV_2 = \lambda_2 BV_2 \end{cases} \Longrightarrow \begin{cases} V_1^TAV_2 = 0 \\ V_1^TBV_2 = 0 \end{cases}. \tag{303}$$

Hence $[V_1, V_2, \dots, V_n]^TA[V_1, V_2, \dots, V_n]$ is a diagonal matrix. The same is valid for matrix B.

1. Coupled oscillators

Consider an oscillating system of two unequal masses.



The equations of motion are

$$\begin{aligned} m_1\ddot{x}_1 &= -2kx_1 + kx_2 \\ m_2\ddot{x}_2 &= kx_1 - 2kx_2, \end{aligned} \tag{304}$$

after introducing vector $u = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, can be presented in a matrix form

$$B\ddot{u} + Au = 0 \tag{305}$$

where

$$B = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad A = k\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}. \tag{306}$$

We are looking for the solution in the form

$$u = u_0 e^{i\omega t}. \tag{307}$$

Hence we obtain equation

$$Au - \omega^2 Bu = 0. \tag{308}$$

Problem F.1 *Choose in the abovementioned problem $m_1 = 1, m_2 = 2, k = 1$. Check up that the solution of Eq. (308) is*

$$\begin{aligned} \omega_{1,2} &= \sqrt{\frac{3 \mp \sqrt{3}}{2}} \\ u_1 &= \begin{bmatrix} \sqrt{3}-1 \\ 1 \end{bmatrix} \\ u_2 &= \begin{bmatrix} \sqrt{3}+1 \\ -1 \end{bmatrix}. \end{aligned}$$

Show, the solution of the problem is

$$\begin{aligned} x_1 &= (\sqrt{3}-1)u_1 + (\sqrt{3}+1)u_2 \\ x_2 &= u_1 - u_2, \end{aligned}$$

where

$$\begin{aligned} u_1 &= a\exp\{i\omega_1 t\} \\ u_2 &= b\exp\{i\omega_2 t\}. \end{aligned}$$